

THIS WEEK

EDITORIALS

WORLD VIEW Throw open your lab books, says Jonathan Schooler **p.437**

SUCKER PUNCH Egg pheromone forces male squid to fight **p.438**

VISAS UK government to give scientists priority under immigration reform **p.440**

Animal rights and wrongs

A Nature survey shows the pernicious impact of activism on biomedical scientists. More institutions must offer researchers the training they need to stand up for their work.

The results of a *Nature* poll of scientists involved in animal research reveal that nearly one-quarter of respondents have been negatively affected by animal-rights activists, or seen it happen to someone they know. In some places, including the United Kingdom, the figure is higher than one-third. The large number of people affected will surprise many of *Nature's* readers. Researchers have suffered fire bombings, physical attacks, destruction of personal property and campaigns of harassment. But the statistics do not necessarily reflect the current prevalence of violent activist behaviour — rather, they reveal how such activity instils a lingering fear that is difficult to forget (see page 452).

The survey shows how corrosive animal-rights extremism can be. It is clear that many of those who perpetrate it remain unrepentant and determined to continue their efforts to terrorize researchers, but there are positive signs. Little more than 15% of poll respondents who were affected by activism said the tactics drove them to change the direction or practice of their research, and several who did make changes said that they mostly became more selective about who they talked to or how they presented their work on the Internet.

There are welcome signs that the tide of violent activity may be turning, especially in the United Kingdom. Several factors could be at work. Tougher legislation might be having an effect; in the past few years, Britain and the United States have both introduced laws that reinforce the seriousness of acts of vandalism intended to bully and blackmail those connected to animal research. Groups in favour of such research have also helped to calm the violence. Pro-Test, an organization based in Oxford, UK, which this week celebrates its fifth anniversary, has managed to counter a campaign of misinformation and intimidation that almost scuttled plans to build a biomedical research facility at the University of Oxford (see page 457). Other groups have begun to follow Pro-Test's lead, including an offshoot at the University of California, Los Angeles, which has been repeatedly targeted by activists. Proactive campaigns and pressure on lawmakers to protect the public's investment in research have aided the backlash against extremism. But these are only part of the solution.

Scientists regularly face the dilemma of how open to be about their animal research. Non-disclosure, even in the scientific literature, is common, according to a recent survey by the UK National Centre for the Replacement, Refinement and Reduction of Animals in Research. Such a lack of openness, it added, could impede reproduction and replication of previous work (C. Kilkenney *et al.* *PLoS ONE* 4, e7824; 2009). Findings such as these have led many journals, including *Nature*, to adopt more-explicit rules about what is to be reported in the literature (see go.nature.com/5zbqp4).

Talking to the public remains crucial. Sometimes, the threat of violence means that individual researchers will not wish to engage directly with the public and should even be cautioned against doing so. But there is no excuse for institutions that house animal

research — including most research universities — not to have vigorous and well-defined programmes to explain what goes on within their walls. Institutions should publicize the high standards that they are required to meet before they can use animals. They should also discuss their strategies to replace animals with more sophisticated

"There is no excuse for institutions not to explain what goes on within their walls."

research tools, refine research practice and reduce the overall number of animals used. If they have no such strategies, institutions should develop them as a priority.

Some scientists who work with animals are already willing to explain the importance of their research. Others should follow their lead. *Nature's* survey found that more than 50% of researchers were encouraged by their institutions to engage with the public, yet not much more than one-quarter felt they were given the necessary training or support. This is unacceptable: the resources are out there, including tips on how to communicate effectively and how best to respond to personal threats.

Activists often attempt to marginalize researchers, isolating them from their institutions and the wider community. If researchers build better and stronger bonds with both, they can ensure that it is the extremists who are marginalized. ■

On the rebound

A critical report has fuelled arguments about the benefits of energy efficiency.

The ghost of William Jevons has haunted energy researchers in recent months, provoking debates on whether our best efforts to use less energy will merely lead us farther down the road of consumption. Jevons, a British economist, suggested in 1865 that increasing energy efficiency could backfire because it would allow further resource exploitation. He was thinking about coal at the time. And, indeed, as people got better at converting this black rock's energy into useful work, the work itself expanded.

The 'Jevons paradox' has persisted ever since. Today, the notion that there could be some modest 'rebound' effect that negates gains in energy efficiency is well known. For example, drivers of fuel-efficient vehicles might be inclined to drive more often, simply because they can afford to do so. It takes energy to create and install energy-efficient equipment; and money saved on energy could be spent elsewhere, so ultimately contributing to economic activity, which drives up energy consumption and greenhouse-gas emissions. (Indeed, one reason for

the strong support for energy-efficiency measures is that they give people more money to spend, not necessarily to save.) According to another rebound effect, if fuel-efficiency regulations for vehicles in the United States and Europe curb petrol consumption, that should suppress the price of oil and encourage its use in other sectors or countries.

Last week, the Breakthrough Institute, based in Oakland, California, released a strongly phrased analysis that argues that these rebound effects are so large they could overwhelm many or all of the gains from energy-efficiency measures (see go.nature.com/m5ky3g). The advocacy group calls for massive investment in the research and development of low-carbon energy to counter what it says are over-optimistic assumptions about what can be accomplished through greater efficiency.

This is controversial stuff, and also frustrating. These are fundamentally interesting questions, but part of a debate that circles on itself. Was it really fair for Jevons to blame coal's expansion and technological diffusion on efficiency? Do people actually save money when their energy bills go down, or do they spend more money? Is the spread of refrigerators around the world purely because they have become more energy efficient, or because they are useful devices that keep food fresh for days on end? Is it the energy bill or the sticker price that people worry about when buying an appliance? And isn't technological

diffusion and energy access throughout the world a good thing?

The rebound effects need to be considered, but they do not have to be viewed as paradoxical: they amount to economic expansion. Indeed, some researchers think that energy efficiency itself is a fundamental driver of economic growth, freeing up resources that can be used for other things, the deployment of low-carbon energy among them.

"The world cannot solve all of its energy and climate woes with energy efficiency alone."

Despite its concerns about the rebound effect, the Breakthrough Institute argues that energy efficiency should nonetheless be pursued for exactly these reasons. Encouragingly, the discussion prompted by its report has led to plans from academics and industry experts on all sides of the debate to

meet to wade through these issues.

The debate indicates that there must be deeper study of what energy efficiency could do if systematically deployed across an entire economy. The world cannot solve all of its energy and climate woes with energy efficiency alone; low-carbon energy technologies must be developed as well. But there seems to be no fundamental physical or economic reason that countries can't decrease their overall energy consumption while maintaining growth, and thus put the ghost of Jevons to rest. ■

Rights for all

Scientists should push for fair treatment of Turkish academics arrested on little evidence.

Earlier this month, a Turkish court acquitted the sociologist Pinar Selek on terrorism charges — the third time she has been cleared of causing a 1998 explosion at an Istanbul spice market that killed seven people. Selek's real crime, in the eyes of the Turkish government, seems to have been contact with Kurdish separatist groups as part of her academic research. Prosecutors have again appealed the not-guilty verdict, to howls of outrage from human-rights groups in Europe, which is watching Turkey's conduct in the case with interest as the nation seeks membership of the European Union (EU).

A less-scrutinized and less-supported cause are the dozen or so Turkish academics arrested — among hundreds of others — as part of a government crackdown on 'deep-state networks'. In the past, these are said to have violently targeted those whom they consider to be Turkey's enemies, including Kurdish separatists. Turkish police say that they have uncovered a tightly controlled, hierarchical organization — called Ergenekon — that conspires to destabilize the government through political violence and pave the way for a military coup.

The government frames its Ergenekon investigation as a step towards democratic reform, but detractors say it has also been used as an excuse to round up and silence government critics, particularly vociferous defenders of secularism in politics and education.

An independent analysis of the investigation by the Istanbul-based analyst Gareth Jenkins says that many of those arrested seem to be linked to one another by little more than their opposition to Islamic conservatism and the ruling Justice and Development Party (AKP). That certainly describes the arrested academics — most of whom hailed from the conservative secular elite — including several university rectors who had decried attempts to lift the headscarf ban in Turkish universities, and criticized, for example, the appointment of AKP supporters to key academic positions, such as in the government's science-funding agency TÜBİTAK.

Most of the arrested academics have been released on bail, but two are still being held: Mehmet Haberal, the US-trained rector of Baskent University in Ankara, who pioneered transplant surgery in Turkey,

and physician Fatih Hilmioğlu, former rector of İnönü University in Malatya.

The two have now been detained for almost two years, and there is no sign that they will be released, or have their trials heard, any time soon. Turkish law allows indicted individuals to be kept in long-term detention only if there is a danger that they might either destroy evidence or flee. Supporters say that both scientists are being treated for poor health and neither danger applies.

Questions have been asked on their behalf. In December 2009, Carol Corillon, executive director of the US National Academies' committee on human rights, expressed concern about the arrest of the academics in *University Values*, an electronic bulletin. And last year, the Turkish Academy of Sciences raised the detentions with Turkish President Abdullah Gül, who said he could not interfere with the independent courts.

Mostly, however, silence rules. The huge Ergenekon trial is moving very slowly, and those out on bail have been warned not to prejudice its outcome by discussing their cases publicly. The investigation has engendered an atmosphere of fear in universities. Academics are keeping their heads down and trying to enjoy the genuine attempts that Turkey is making to raise its science base towards EU norms. TÜBİTAK is also keeping quiet, afraid of further charges of political interference in scientific matters, such as it attracted in 2009, when it pulled the Darwin-celebratory cover of its popular-science magazine *Bilim ve Teknik* and demoted its editor (see *Nature* **458**, 259; 2009).

Meanwhile, most human-rights groups have proved less keen to jump to the defence of those on the political right than those on the left, including Selek. Jenkins' report did stress that the Ergenekon investigation successfully identified some deep-state operatives, probably involved in political violence. And phone-tap transcripts have exposed unpleasant and antidemocratic right-wing sentiment among some of the indicted individuals.

But according to the European Convention on Human Rights, everyone has the right to a fair trial in a reasonable time. The Turkish government needs to feel the world's eyes on the actions of its judiciary. And the international scientific community should spread the word on the indicted and detained academics, and monitor how fairly their cases are handled. The last thing Turkish scientists need

right now — as research investment is growing but the legal and administrative framework is not always keeping up — is to be afraid to criticize how their promising scientific landscape is developing. ■

➔ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhunq



Unpublished results hide the decline effect

Some effects diminish when tests are repeated. **Jonathan Schooler** says being open about findings that don't make the scientific record could reveal why.

Many scientifically discovered effects published in the literature seem to diminish with time. Dubbed the decline effect, this puzzling anomaly was first discovered in the 1930s in research into parapsychology, in which the statistical significance of purported evidence for psychic ability declined as studies were repeated. It has since been reported in a string of fields — both in individual labs (including my own) and in meta-analyses of findings in biology and medicine. The issue has been recognized in some circles within the scientific community, but rose to wider prominence last December when it was discussed in an article in the magazine *The New Yorker*.

Some scientists attribute the decline effect to statistical self-correction of initially exaggerated outcomes, also known as regression to the mean. But we cannot be sure of this interpretation, or even test it, because we do not generally have access to 'negative results': experimental outcomes that were not noteworthy or consistent enough to pass peer review and be published.

How could the availability of unpublished results be improved? I suggest an open-access repository for all research findings, which would let scientists log their hypotheses and methodologies before an experiment, and their results afterwards, regardless of outcome. Such a database would reveal how published studies fit into the larger set of conducted studies, and would help to answer many questions about the decline effect.

Availability of unpublished findings could also address other shortcomings of the current scientific process, including the regular failure of scientists to report experiments, conditions or observations that are inconsistent with hypotheses; the addition or removal of participants and variables to generate statistical significance; and the probable existence of numerous published findings whose non-replicability is shrouded because it is difficult to report null results.

To address the decline effect, such a database could pinpoint whether the phenomenon reflects how scientists design experiments, how they write them up or how journals decide what to publish. It could be used to explore whether genuine changes in studied phenomena could stem from conventional mechanisms; for example, in social sciences, decline effects could be the result of participants no longer being naive about the effect under investigation. Less likely, but not inconceivable, is an effect stemming from some unconventional process. Perhaps, just as the act of observation has been suggested to affect quantum measurements, scientific observation could subtly change some scientific effects. Although the laws of reality are usually understood to be immutable, some physicists, including Paul Davies, director of the BEYOND: Center for Fundamental Concepts in

Science at Arizona State University in Tempe, have observed that this should be considered an assumption, not a foregone conclusion.

More prosaic explanations for the decline effect include the previously mentioned regression to the mean. If early results are most likely to be reported when errors combine to magnify the apparent effect, then published studies will show systematic bias towards initially exaggerated findings, which are subsequently statistically self-corrected (although this would not account for the typically linear nature of the decline).

Publication bias could also be responsible. Researchers might only be able to publish initial findings on an effect when it is especially large, whereas follow-up studies might be more able to report smaller effects. Other potential answers include unreported aspects of methods, exclusive reporting of findings consistent with hypotheses, changes in

researcher enthusiasm, more rigorous methodologies used in later studies, measurement error resulting from experimenter bias and the general difficulty of publishing failures of replication.

An open-access database of research methods and published and unpublished findings would go a long way towards testing these ideas. For example, both the regression to the mean and degradation of procedure explanations assume that early published studies benefit from being at one statistical end of a larger body of (unpublished) findings. Publication bias and selective reporting of data are similarly difficult to investigate without knowing about unpublished data.

An open-access repository of findings would be difficult to introduce. It would need an automated protocol to enable study methods and results to be entered and retrieved. Some way to assess the quality of the work would be required — perhaps

through open-access commentaries moderated in a manner similar to Wikipedia. We would need to assure the qualifications of researchers who use it, and maintain a blackout period to protect hypotheses and findings prior to publication. Reluctant scientists would need incentives — and perhaps new rules from funders — to take part.

Such challenges would not be insurmountable. Similar, if more narrowly defined, databases have already been set up for clinical trials (<http://clinicaltrials.gov>) and educational research (<http://pslcdatashop.web.cmu.edu>). A good starting point might be to develop a host of subject-specific repositories. However it is implemented, we need a better record of unpublished research before we can know how well the current scientific process, based on peer review and experimental replication, succeeds in distinguishing grounded truth from unwarranted fallacy. ■

Jonathan Schooler is a professor of psychology at the University of California, Santa Barbara.
e-mail: schooler@psych.ucsb.edu

WE NEED A BETTER
RECORD
TO LEARN HOW
WELL SCIENCE
DISTINGUISHES
TRUTH
FROM
FALLACY.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/xbvqit

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

EXOPLANETS

Earth-sized, not Earth-like

NASA's Kepler mission to find habitable planets orbiting Sun-like stars has turned up its first rocky planet. The project uses the Kepler space telescope to identify extrasolar planets by watching for dips in the intensity of light from up to 170,000 target stars.

Natalie Batalha of San Jose State University in California and her group spotted Kepler 10b, which is about 4.56 times the mass of Earth. Although similar in size to Earth, its orbit lasts just 0.84 days, making it likely that the planet is a scorched, waterless world with a sea of lava on its starlit side.

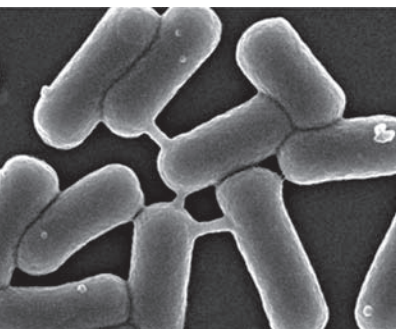
Astrophys. J. 729, 27 (2011)

MICROBIOLOGY

Molecular mail by nanotube

Bacteria can communicate by sending molecules to each other along nanotube bridges. Researchers previously thought that bacteria exchanged molecules primarily by secreting and sensing them.

Gyanendra Dubey and Sigal Ben-Yehuda at the Hebrew University of Jerusalem in Israel observed tubular protrusions between neighbouring *Bacillus subtilis* cells (**pictured**) and showed that they are used to exchange



ELSEVIER

molecules and proteins. The nanotubes, which are typical of multicellular organisms, also form between *B. subtilis* and other bacterial species.

The bridging mechanism can transfer features such as antibiotic resistance between cells, the authors say.

Cell 144, 590–600 (2011)

CANCER

Cyclin through drug resistance

Breast-cancer tumours can become resistant to therapy if they express abnormally high levels of the gene for a protein called cyclin E. The finding

suggests that a combination of drugs targeted at specific molecules may one day be appropriate for such tumours.

The therapeutic antibody trastuzumab is designed for breast cancers that overexpress the gene *HER2*, but patients frequently become resistant to the treatment. José Baselga, now at Massachusetts General Hospital in Boston, and his colleagues analysed tumour cells from 34 patients treated with trastuzumab. They found that the median time it took for tumour progression to recommence was cut by more than half in those whose tumours overexpressed cyclin E.

CHEMISTRY

One catalyst, two products

Many chemical reactions produce two enantiomers — molecules that are mirror images of each other — but



R. HANLON

ZOOLOGY

Fighters egged on by pheromones

Male squid are propelled into an immediate fighting frenzy by contact with squid eggs — a response, researchers say, to a pheromone identified on the eggs' surface. The protein, *Loligo* β -MSP, is the first aggression pheromone from a marine animal to be characterized at the molecular level. Its similarity to proteins found in mammalian seminal fluids suggests that these may also have a role in sexual competition. Roger Hanlon at the Marine Biological Laboratory in Woods Hole, Massachusetts,

and his team isolated the protein and applied it to the surface of a clear glass flask containing eggs. They found that male squid (*Loligo pealeii*, **pictured**) began to push and bite each other violently within seconds of touching the glass, competing for females as they do after contact with natural eggs. The pheromone, which is made by female reproductive glands, may help to focus males' competitive aggression when mature, receptive females are nearby.

Curr. Biol. doi:10.1016/j.cub.2011.01.038 (2011)

Cyclin E acts with another protein called CDK2 to usher cells through a specific phase in the cell cycle. A compound that inhibits CDK2 slowed the growth of trastuzumab-resistant tumour cells transplanted into mice.

Proc. Natl Acad. Sci. USA
doi:10.1073/pnas.1014835108 (2011)

often only one is desired. Researchers have developed a catalyst that changes its configuration when exposed to light and that can generate one enantiomer or the other, depending on the position of the catalyst's rotating parts.

The catalyst is based on a molecular 'motor' comprising two parts rotating around a fixed axle. The motor moves through a four-step cycle in which each change in position is triggered by light. At each position, the relative orientation of the two moving parts controls the spatial position of the reactant in a chemical reaction, producing either one or the other enantiomer, or a mix of the two.

Using just one molecule of the catalyst, Jiaobing Wang and Ben Feringa at the University of Groningen in the Netherlands produced both enantiomers in a sequential manner. *Science* doi:10.1126/science.1199844 (2011)

IMMUNOLOGY

Modulating malaria's mayhem

The malaria parasite wreaks havoc in the body, in part by activating receptors that normally detect the presence of microbes and initiate innate immune responses. By interfering with these 'Toll-like' receptors in a mouse model of cerebral malaria, the worst symptoms of the disease can be prevented, say Ricardo Gazzinelli at the University of Massachusetts Medical School in Worcester and his team.

By activating Toll-like receptors, the malaria parasite triggers the excessive release of inflammatory immune molecules called cytokines. The researchers used a small molecule to block Toll-like-receptor activation in mice after they had been infected with malaria. The inhibitor vastly improved the animals' survival and decreased the prevalence of cerebral malaria symptoms such as seizures, even though it did not decrease the number of

parasites. The strategy might help to cut malaria mortality in the absence of a preventive vaccine, the authors say.

Proc. Natl Acad. Sci. USA
doi:10.1073/pnas.1015406108 (2011)

HYDROLOGY

Melting to the max

As the climate warms, a larger proportion of precipitation is likely to take the form of rain and less as snow. The snow that does still accumulate on mountains during winter will subsequently melt faster in the spring, increasing peak discharge and flood risk.

Amilcare Porporato and his colleagues at Duke University in Durham, North Carolina, used a model to calculate how seasonal temperature affects snow accumulation, melting and mountain run-off. In the model, peak discharge increases to a maximum in a warming environment despite reduced snowpack, and then drops off as less and less snow accumulates in winter.

In the real world, climate-change-driven alterations in the timing and volume of mountain snowmelt could increase the risk of both flooding and summer droughts in the catchment areas of meltwater-fed rivers. *Geophys. Res. Lett.* doi:10.1029/2010GL046477 (2011)

MOLECULAR BIOLOGY

When DNA repair goes wrong

Mistakes in the DNA sequence are constantly being repaired, but some repair mechanisms are themselves error-prone. One, called break-induced replication (BIR), is initiated at sites at which both DNA strands are broken, and mimics the process of DNA replication. However, its rate of frameshift mutation — a type of mutation that almost always results in loss of gene function — is up to 2,800 times higher than that of normal replication.

COMMUNITY CHOICE

The most viewed papers in science

ENVIRONMENTAL CHEMISTRY

Tobacco smoke, take three

HIGHLY READ
on pubs.acs.org
in January 2011

The dangers of second-hand cigarette smoke are well known, but nicotine can linger on indoor surfaces and generate 'third-hand' smoke, either through its re-emission into the air or by direct ingestion from surfaces. Researchers have found that nicotine on surfaces can produce aerosols of potentially toxic oxidation products, by reacting with certain oxidizing gases when exposed to conditions typical of those found in the home.

Yael Dubowski and her team at the Israel Institute of Technology in Haifa doped cellulose powder, paper and cotton samples with a nicotine suspension and then exposed them to ozone and nitrogen oxide gases. Aerosols were collected and analysed.

The authors found that relative humidity and the type of surface affected the number and size distribution of the aerosol particles. Aerosol formation was greatest on cellulose in dry air, whereas much fewer aerosols formed from paper. *Environ. Sci. Technol.* 45, 328–333 (2011)

Anna Malkova at Indiana University-Purdue University Indianapolis and her team measured mutation rates during BIR in yeast. They found that errors occurred not just close to the site of the double-stranded break, but also farther down the DNA. Most of the mistakes were generated by DNA polymerase enzymes, which replicate DNA and are generally thought to be accurate. The authors say that one of the factors leading to the high mutation rate was elevated levels of nucleotides and that BIR could be an important source of the mutations that drive cancer and evolution. *PLoS Biol.* 9, e1000594 (2011)

PALAEOANTHROPOLOGY

No marathons for Neanderthals

Fossilized heel bones suggest that Neanderthals were not built to run long distances.

By observing contemporary human distance runners on a treadmill, David Raichlen at the University of Arizona in Tucson and his team found that athletes with shorter heel



bones, or calcanei (pictured, white line), tended to run more economically — consuming less oxygen to maintain the same speed — than those with longer ones. Calcanei from 7 Neanderthals were, on average, longer than the same bone from 13 ancient humans and the 8 treadmill runners.

Our African-dwelling ancient ancestors may have evolved feet adapted to chasing animals to the point of heat exhaustion. Neanderthals, who lived in colder, forested European and Asian habitats, may instead have attacked prey at close range, the authors say. *J. Hum. Evol.* 60, 299–308 (2011)

NATURE.COM

For the latest research published by Nature visit:
www.nature.com/latestresearch

D. RAICHLEN

SEVEN DAYS

The news in brief

POLICY

US budget blitzed

The US House of Representatives slashed billions of dollars from government science and environmental-protection agencies in a bill passed on 19 February. With the current federal budget expiring on 4 March, the bill was a 'continuing resolution' to fund the government until the 2011 fiscal year ends in September. The Senate is likely to reject the bill when it returns from recess on 28 February, but the landscape of cuts offers clues to Republican targets in the budget battle ahead. "It is truly, truly frightening," says Sally Yozell, director of policy at the National Oceanic and Atmospheric Administration in Washington DC. See go.nature.com/jwwkwf for more.

Brazil cuts budget

Brazil's new president, Dilma Rousseff, is cutting the country's 2011 science budget as part of a broad strategy to lower public spending and keep inflation at bay. The Ministry of Science and Technology will receive only 6.4 billion real (US\$3.84 billion), it emerged last week, down from the 7.4-billion-real plan agreed by Congress. See go.nature.com/tpkeb7 for more.

Food safety at risk

The broad food-safety law signed by US President Barack Obama last month still leaves food-safety oversight fragmented across federal agencies, according to a report released on 16 February by the US Government Accountability Office. The biennial report, which catalogues government programmes at "high risk" for waste, fraud or abuse, also raises concerns about the Food and Drug Administration's

oversight of drugs and medical devices produced abroad, and the Environmental Protection Agency's monitoring of toxic chemicals. Similar issues were flagged in the 2009 edition of the report.

UK visa rules

Scientists needing visas to work in the United Kingdom will get priority over other workers under an immigration quota that takes effect from April. The Campaign for Science and Engineering, an advocacy group based in London, had feared that the government's decision last November to limit immigration of highly skilled workers from outside the European Union would put young, low-salaried researchers at a disadvantage. But the points system used to rank applications, published

on 16 February, shows that the government has decided to score scientists highly. The impact of related plans to limit student immigration is not yet clear. See go.nature.com/mmg98w for more.

BUSINESS

Huge pharma deal

Drug-maker sanofi-aventis has struck a deal to buy biotechnology company Genzyme of Cambridge, Massachusetts, ending more than six months of negotiations, it announced on 16 February. Sanofi, based in Paris, will pay US\$20.1 billion (\$74 per share), but the deal includes extra payments tied to the success of Genzyme's drug pipeline. Sanofi initially offered \$69 per share, which Genzyme rejected in August 2010. See page 449 for more.

RESEARCH

Egypt's antiquities

On 20 February, Egypt reopened its archaeological sites and six of its antiquities museums, including Cairo's Egyptian Museum, to visitors. Zahi Hawass, head of Egypt's Supreme Council of Antiquities in Cairo, said that, despite some looting during the revolution that deposed former president Hosni Mubarak, "relatively few" objects were missing from the museum.

Anthrax report

Scientific evidence used by the Federal Bureau of Investigation (FBI) cannot alone prove the origin of anthrax bacteria mailed in the United States in 2001, which killed 5 people and made 17 ill, says a National Research



M. HUNTER/GETTY IMAGES

Earthquake devastates Christchurch

The 6.3-magnitude earthquake that struck Christchurch, New Zealand, on 22 February was probably part of a sequence of aftershocks that followed a 7.0-magnitude earthquake on 3 September 2010, according to the US Geological Survey. Although it was more

powerful, September's quake caused no deaths. The aftershock was shallower, much closer to Christchurch, and struck at a busy hour, just before 1 p.m. local time; 65 deaths were confirmed as *Nature* went to press. See go.nature.com/5cezck for more.

Council committee. The panel's independent review, released on 15 February, questioned FBI conclusions about the laboratory flask that was pinpointed as the source of the *Bacillus anthracis* bacteria, and about the skills required to grow anthrax spores. The FBI responded that it was correct to accuse the late microbiologist Bruce Ivins of the attacks, taking into account all the evidence in the case. See go.nature.com/21sfju for more.

Green economy

Investing US\$1.3 trillion — or 2% of global gross domestic product — every year between now and 2050 in environmental projects will spur economic growth, according to a report by the United Nations Environment Programme. The study, released in Nairobi on 21 February, recommends investment in sustainable agriculture, low-carbon energy and natural resources such as fisheries, drinking water and forests. The report is a primer for a UN conference on sustainable development, to be held in Rio de Janeiro, Brazil, in May 2012. See go.nature.com/ae5uzu for more.

Whaling halted

Japan has called off its whale hunt in Antarctic waters about a month before the end of the whaling season. The

decision follows sustained attacks by the anti-whaling group the Sea Shepherd Conservation Society, based in Friday Harbor, Washington. The society has sparred with Japanese fishing vessels every year since 2005 to disrupt their hunting in the Southern Ocean. This year's early halt, announced on 18 February, is a first. See go.nature.com/xukxhx for more.

Satellite switch-off

NASA turned off its Wide-field Infrared Survey Explorer (WISE) as scheduled on 17 February. The telescope, launched in December 2009, completed its ten-month mission to record the entire sky in infrared wavelengths by October, and extended its mission by cataloguing near-Earth asteroids. The mission releases the first 57% of its data in April, and its full data set will be available in 2012. See go.nature.com/fy1tg8 for more.

Computer wins quiz

In an internationally publicized victory for computer science, the supercomputer Watson trounced two human contestants on the US television quiz programme *Jeopardy!* last week. The system is based on IBM's DeepQA project, one of a number of projects that aim to analyse and answer 'natural-language' questions posed in

standard English. In *Jeopardy!*, this includes puns, quips and hidden ambiguities. Similar systems could create more powerful search engines; for scientists, this might make it easier to trawl through research literature. See go.nature.com/u783dz for more.

PEOPLE



Science chief quits

Australia's chief scientist, Penny Sackett (pictured), has tendered her resignation halfway through her five-year term. She will step down on 4 March. Sackett, formerly director of astronomy and astrophysics research at the Australian National University in Canberra, cited "both personal and professional reasons" in her surprise announcement on 18 February. No further explanations had emerged as *Nature* went to press.

COMING UP

25 FEBRUARY

Elections in the Republic of Ireland are expected to oust the long-standing incumbent government, Fianna Fáil, which has heavily supported science for the past two decades.

28 FEB–1 MARCH

The Royal Society in London holds a meeting on how to reduce agricultural greenhouse-gas emissions.

go.nature.com/ffugs9

Scripps president

Biochemist Michael Marletta will be the next president of the Scripps Research Institute, succeeding chemist Richard Lerner who has held the position for the past 25 years. The institute, which is located in La Jolla, California, announced the change on 20 February. Currently based at the University of California, Berkeley, Marletta will move to Scripps in July and take the helm there from 1 January 2012.

Yamada steps down

Tachi Yamada is to retire after five years as president of the Global Health Programme at the Bill & Melinda Gates Foundation in Seattle, Washington, the foundation announced on 14 February. Before joining the foundation, Yamada, who will step down in June, was head of research and development at London-based drug giant GlaxoSmithKline. One of his focuses at the foundation has been to enlist the support of pharmaceutical companies in developing drugs and vaccines for the developing world.

NATURE.COM

For daily news updates see:

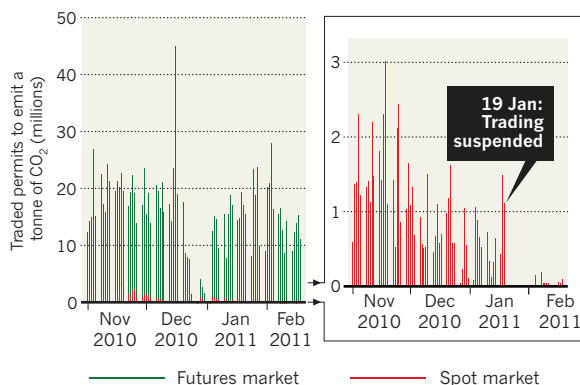
www.nature.com/news

TREND WATCH

The world's largest carbon-trading market is still reeling from a 19 January shut-down after €30 million (US\$41 million) in carbon credits was stolen from insecure accounts. The European Commission had to suspend immediate trading of allowances to emit carbon dioxide; only seven of the European market's 30 national registries have since opened up allowances for trading. Total traded volumes have not been dented, however, because most of the trading is based on the future price of allowances.

THEFT ON THE CARBON MARKET

Although overall trading volumes have been mostly unaffected by a suspension forced on Europe's carbon market, immediate (spot) trading only resumed on 4 February.



NEWS IN FOCUS

RESEARCH European centre aims to turn sea creatures into lab workhorses **p.444**

BUSINESS When big pharma swallows a biotech, can innovation survive? **p.449**

Q&A A veteran of US climate policy reflects on battles still to come **p.451**

ANIMAL RESEARCH *Nature* survey reveals widespread harassment. **p.452**



Iranian President Mahmoud Ahmadinejad (centre) visiting the Natanz nuclear enrichment facility.

REUTERS/IRAN'S OFFICIAL PRESIDENTIAL WEBSITE

NON-PROLIFERATION

Clash over Iran's capability

Effects of sanctions and computer worm on uranium production are disputed.

BY SHARON WEINBERGER

Experts at two prominent organizations are clashing over whether Iran is improving its uranium enrichment capability, a key measure of its ability to produce a nuclear weapon.

On one side is David Albright, president of the Institute for Science and International Security (ISIS) in Washington DC, and a widely quoted authority on Iran's nuclear programme. In a report (go.nature.com/5rgzpa) released by ISIS on 15 February, Albright and his co-authors find that the Stuxnet computer

worm — a computer program that hit many industrial systems in 2009 and 2010, and is widely thought to have been targeted at Iran's nuclear programme — probably destroyed about 1,000 centrifuges located at a fuel enrichment plant at Natanz, in central Iran. "It set them back," Albright told US television personality Stephen Colbert in an interview, echoing the US government's optimism.

Just a few weeks before, however, an analysis (go.nature.com/sihqyl) published by Ivanka Barzashka, a physicist then working at the Federation of American Scientists (FAS), also in Washington DC, came to a very different

conclusion. It is impossible to estimate the effect of Stuxnet on Iran's nuclear programme, and more to the point, Iran's centrifuges are, on average, performing much better than the previous year, the report says. The ISIS assumption about Stuxnet destroying centrifuges "can lead to dangerous conclusions, such as that we are slowing down Iran's programme, and that gives you a sense of complacency", says Barzashka, now a visiting scholar at the Bulgarian Academy of Sciences, Sofia.

The technical debate between the FAS and ISIS centres on the effectiveness of Iran's IR-1 centrifuges, the fast spinning machines that ►

► separate uranium-235 from the heavier uranium-238 isotope. A number of outside organizations, including the FAS and ISIS, have focused on how efficiently those centrifuges can enrich uranium, measured as kilograms of enriched uranium produced per standard 'separative work unit' (kg SWU). This number would help indicate how quickly Iran, if it so chose, could enrich enough material for a nuclear weapon.

Barzashka bases her derivation of SWU on data provided by the International Atomic Energy Agency (IAEA), which conducts regular inspections of Iran's nuclear facilities. On the basis of those numbers, she estimates that Iran's ability to enrich uranium is increasing. This suggests that Stuxnet has not been as detrimental as Albright suggests, and that Iran has actually improved its overall efficiency with centrifuges.

The FAS entered the debate in 2009 (see 'Spinning the numbers') when Barzashka and Ivan Oelrich, a retired senior fellow at the FAS, using publicly available IAEA numbers, calculated Iran's centrifuge performance as around 0.44 kg SWU per year. Using the same methods, Barzashka's latest analysis finds 0.77 kg SWU per year. Both numbers are lower than those of ISIS, and lower than many other independent estimates. But Barzashka maintains that just as others were overestimating Iran's capabilities in the past, they are now perhaps overestimating the effect of the Stuxnet worm and ignoring evidence that Iran's capabilities are actually improving.

SPINNING THE NUMBERS

Estimates of Iran's pace of uranium enrichment have varied over time. Figures seem to converge but there is uncertainty about the facts behind the trend.

	kg SWU/yr	Source
16 MAY 2006	1.46	Jeffrey Lewis go.nature.com/izqsmj
1 NOV 2007	2	David Albright go.nature.com/ghrzfv
17 JAN 2008	1.36	Richard Garwin go.nature.com/4psvte
27 FEB 2009	2.2	go.nature.com/ghrzfv
22 SEP 2009	2.1	Ali Akbar Salehi go.nature.com/wp4kxs
25 SEP 2009	0.44	Ivan Oelrich & Ivanka Barzashka go.nature.com/vnuhzz
21 JAN 2011	0.77	Ivanka Barzashka go.nature.com/sihqyl
15 FEB 2011	0.9	David Albright et al. go.nature.com/5rgzpa

Albright now places Iran's enrichment capacity at 0.9 kg SWU per year, partly on the basis of government information that is not available to other experts. This is a downward trend relative to earlier numbers from ISIS, but despite the fact the two groups' numbers seem to be converging, the conclusions that they draw from those numbers have continued to fuel an acrimonious debate.

In December 2010, Albright e-mailed the head of the FAS, calling Oelrich and

Barzashka's Iran estimates a "bogus claim". Oelrich and Barzashka — their departures from the FAS were unrelated to the matter — disagree. "Albright doesn't assume any value, he bases his value on expert opinion," says Oelrich. "And we don't assume any value; we calculate it based on IAEA results."

Albright says that if Iran were to pursue a nuclear weapon at full speed, it would take the country only about six months to enrich a sufficient quantity of uranium at Natanz. The Obama administration, according to Albright, uses a more conservative estimate of 10–14 months. Barzashka's latest estimate of five months is in close agreement with Albright.

Barzashka says that, in such a 'breakout scenario' of full-speed production, underestimating Iran's progress "takes away from the urgency of a diplomatic breakthrough".

But for some, the absolute numbers matter less than the debate itself. "The one thing we've learned from the Iraq debacle is that you have to have technical debates; everyone's assumptions have to be challenged," says Jacqueline Shire, a member of the United Nations Panel of Experts, which looks at the effectiveness of sanctions against Iran.

Shire, who used to work at ISIS, warns that getting "bogged down in SWU numbers" can also divert attention from the real issue. "One can very easily start to get buried in minutiae and lose sight of the big picture," she says, "which is to stop Iran from enriching uranium, and think creatively about how to do that." ■

RESEARCH

Marine biology network launches into choppy waters

Ambitious European project hopes to navigate uncertain funding future.

BY NICOLA NOSENGO

Sometimes good ideas take a while to be picked up. In 1872, Anton Dohrn, a pioneering German biologist, wrote a commentary in *Nature* proposing the foundation of "a net of scientific stations" along European coasts, focusing on marine biology (*Nature* 5, 277–280; 1872). Almost 140 years later, an institute that bears Dohrn's name is leading a twenty-first-century realization of his idea.

The European Marine Biological Resource Centre (EMBRC) will launch this week at a meeting in Naples, Italy, with the Anton Dohrn Zoological Station in Naples (SZN) taking the lead. Linking 15 existing research centres in 8 countries (see 'Marine network'), the project will create an overarching organization for

European research on marine biology, and provide model organisms for studying fundamental molecular biology and for screening drug candidates, for example. But the project has yet to secure the ambitious budget needed to realize its full potential.

Its goal is to make experiments on marine organisms as easy and common as those on mice or fruitflies. "Most of the molecular biology we know today comes from terrestrial species," says SZN president Roberto di Lauro. "Genetics has mostly focused on *Drosophila* or the mouse because they are very easy to grow in a lab." Yet, he points out, most of the world's genetic diversity is found at sea, and some marine species are becoming useful model organisms. The sea squirt *Ciona intestinalis*, for example, is one of the closest invertebrate

relatives of humans and has had its genome sequenced, making it useful for a range of studies. Interest in marine organisms such as *Ciona* is growing rapidly outside the community of marine ecologists.

The EMBRC aims to select a few species as the best candidates to become model organisms and develop technologies to grow them in artificial seawater, and to study and modify their genes. These organisms will be made available to other institutes or companies, along with the technology and support needed to study them. The centre could also produce strains of genetically modified organisms on request for institutes that want

► **NATURE.COM**
To read Anton Dohrn's original vision, see:
go.nature.com/Wel8cN



The sea squirt could become a top model organism at Europe's new marine biology centre.

to grow them on their own premises, as now happens with rodents.

In the preparatory phase, lasting 3 years, €3.9 million (US\$5.3 million) from the European Commission will be used to plan the organizational structure of the new network. "Most of the work will be about managing data," says Detlev Arendt, a developmental biologist at the European Molecular Biology Laboratory in Heidelberg, Germany, who will lead the design of the EMBRC's bioinformatics infrastructure.

The next stage will see partner institutes spend €100 million over 5 years to build up their research capacity. Once the centre is operational, each country will pay for the costs of its own institutes, and will also contribute to the budget of the coordinating body.

This is where some problems might arise. Six other research-infrastructure networks launched by the European Commission in 2006 are now at the end of their preparatory phase, and some are having trouble getting the funds needed to make further progress. "Things are going slower than we expected," admits Jean-Emmanuel Faure of the European Commission's Directorate General for Research. "The timing is bad, with the economic crisis hitting some countries very hard."

Ireland, for example, has withdrawn its financial support to the European Clinical Research Infrastructures Network, based in Paris. And the European Commission has extended the preparatory phase of INSTRUCT, a similar initiative in structural biology, giving it two more years to look for alternative funding sources.

If it can secure funding, the EMBRC has the potential to become a keystone in a global network of marine centres, according to Gary Borisy, director of the Marine Biological Laboratory in Woods Hole, Massachusetts. "We have something similar in the United States with the National Association of Marine Laboratories, but we are not that advanced in terms of standardizing technologies," he says. Borisy hopes that the EMBRC will create opportunities for collaborations between Europe, the United States and Japan, which is also very active in marine genomics. "The first thing to do would be to agree on a few model species and make them worldwide standards," he says.

Making the EMBRC part of a global network is the ultimate long-term goal, agrees di Lauro. The challenge, as Dohrn noted 140 years ago, is that "governments will not easily be induced to sacrifice much money for the progress of this science". ■

MARINE NETWORK

The European Marine Biological Resource Centre will help more than a dozen partner institutions to coordinate their research efforts.



EDUCATION

Faculty members in conflict with president of Japanese university

Allegations over research results highlight the difficulties of investigating high-ranking administrators.

BY DAVID CYRANOSKI

Tension has been high at one of Japan's top research universities. A number of faculty members at Tohoku University in Sendai have called for a fuller investigation into four papers by the university's president, materials scientist Akihisa Inoue, and have expressed dissatisfaction with the university's handling of the matter. There is no reason to conclude that there was any kind of manipulation of data by Inoue, and the accusations over his work are unproven. However, the spat underscores the challenges involved when a university investigates one of its own — especially when the individual under scrutiny is the university's president — as well highlighting the difficulty of verifying experimental results that require extreme precision.

The controversy has been simmering for more than three years, during which critics have repeatedly questioned the results of the four papers¹⁻⁴ by Inoue. An internal committee at the university assessed the criticisms and ruled that a formal investigation was not warranted. In the committee's December 2007 report, senior officials questioned whether the criticism was motivated by "malice" and "divorced from a pure concern for academic development". Since then, university faculty members have repeated the criticisms and raised others. But Inoue, a prolific specialist in an unusual form of alloy called metallic glasses, told *Nature* that his team has unique skills and experience in producing the alloys, which could explain why other scientists have failed to reproduce some of his lab's results.

Collaborators, competitors and critics all recognize Inoue as a pioneer. One colleague says his contribution to the science of bulk metallic glasses is "legion". In 2009, Inoue and William Johnson, of the California Institute of Technology in Pasadena, shared the American Physical Society's prestigious James C. McGroddy Prize for New Materials for their work in the field.

Metallic glasses are formed by cooling a molten alloy so quickly that it doesn't have time to crystallize. This leaves its atoms in a state of disorder — disrupting the regular crystalline structure seen in normal metal alloys. The resulting materials are stronger, more elastic and more resistant to corrosion.



Akihisa Inoue is a leading researcher in bulk metallic glass alloys.

Although broad commercial applications have yet to take off, partly because of the need for novel processing methods, scientists are starting to take advantage of the materials' unique properties to make golf clubs, mobile-phone casings and biomedical devices^{5,6}. "It is one of the most important areas in materials science from a basic scientific point of view, as well as for application-driven developments," says Hans-Jörg Fecht, a materials scientist at Ulm University in Germany.

Inoue has published more than 2,500 articles since 1973. He and his group have reported making numerous glassy alloys and producing samples much thicker than those reported by other groups.

QUESTIONS RAISED

But in May 2007, a series of anonymous letters began arriving at Tohoku University and other places alleging that the four papers¹⁻⁴ co-authored by Inoue in the 1990s contained inconsistencies in the way that the data were presented. The letters also alleged that others in the field had been unable to reproduce the results.

In response to these allegations, Tetsuo Shoji, the university's executive vice-president for research affairs at the time, formed a five-person committee to decide whether a full-scale investigation was warranted. In December

2007, the committee issued a 12-page report that said there were "no rational grounds" for a full investigation. On the problem of irreproducibility, the report said: "various factors including the purity of materials... the cooling method, the protocols, temperature control, moisture control and time control can bring about differences in results, making it easy to imagine how problems in reproducing work might exist among researchers."

Nine metallic-glass experts outside Japan contacted by *Nature* generally lauded Inoue's contribution to the field. At least one, however, had specifically tried to produce some of the metallic glasses described in the papers¹⁻⁴ under discussion but had failed to achieve the large dimensions reported by Inoue. Two others were unable to reproduce other metallic-glass results from Inoue's laboratory.

Inoue disagrees that any difficulties experienced by other groups trying to reproduce his experiments mean that his work is problematic. "We are far ahead of other groups," he says. "If they have the proper level of technological know-how, they should be able to reproduce the work." He adds that his work has been independently reproduced by other groups, although the examples he cites⁷⁻¹⁰ involve significantly smaller sample diameters or glasses of a different composition from those in the four 1990s papers.

Inoue also refers to a paper¹¹ that his group published in 2007 as evidence that the earlier results can be reproduced. This paper reported the production of a zirconium-based metallic glass rod 30 millimetres in diameter. Tsuyoshi Kajitani, a researcher at Tohoku University, has questioned whether the evidence provided in the 2007 paper proved the glassy nature of the sample¹², but Inoue and his co-authors stand by the paper, and Inoue has defended it in writing¹³. Moreover, in response to queries from *Nature*, Alain Yavari, a director of research at the CNRS in France, says that X-ray micro-profiling tests run at the European Synchrotron Radiation Facility in Grenoble as part of another experiment show that a 30-millimetre-thick sample of the zirconium-based material received from Inoue was indeed glassy, supporting the results of the 2007 paper.

Resolution of the broader dispute has been complicated by the fact that Inoue's co-author Tao Zhang has been unable to produce the original laboratory notebooks or glass samples from the experiments that they did for the four disputed papers¹⁻⁴ in the 1990s. Zhang says that the notes and samples accidentally sank in Tianjin Bay in a shipping container when he was moving back to China in 2003. Zhang, now at the Beijing University of Aeronautics and Astronautics, says that he stands by the work.

FURTHER CHALLENGES

Various faculty members at Tohoku University have since challenged both the university's decision not to investigate further and the composition of the initial investigating committee. The conflict has come to focus more on how the university should handle such allegations than on the papers themselves.

In a letter sent to the university's research-affairs committee in February 2008, economics professor Izumi Omura, writing with four other university faculty members, pointed out that the five-person investigating committee did not feature a metallic-glass specialist, and that the specialist brought in to review the technical aspects of the report and "to ensure objectivity" had previously co-authored other papers with Inoue. "That was ludicrous," says Omura.

Shoji, the committee head, defends his decision. "In picking the independent expert, I don't think you necessarily decide based on external appearances," he says. "Quality is what's important."

Omura and other critics also point out, however, that the committee included three people (Shoji, Makoto Watanabe and Keiichi Noe) who had been promoted to their current positions as executive vice-presidents by Inoue since he became university president in 2006. Inoue denies any influence over the committee's make-up. Shoji says that he and other committee members discussed the problem of conflict of interest but decided they would "be able to fairly evaluate the situation". Some critics believe that actual fairness is not enough, and that such a committee must also be seen to

"We see the president as just another scientist and so we leave it up to the university to do any investigation."

Multidisciplinary Research for Advanced Materials sent a petition to the university calling for further explanation of the investigating committee's initial report. Attached to the petition was a covering letter by the institute's director, Fumio Saito (who had not signed the petition himself). Then, just over a month later, Saito sent a letter to all university divisions apologizing for implying in his covering letter that the petition expressed "the collective opinion of the institute". Saito's apology added to the controversy.

In yet another twist, in April 2009 a university committee chaired by Inoue temporarily postponed the granting of emeritus status to two retiring faculty members who had been involved with Omura on a website that collects information about the dispute. Emeritus status is customarily granted to retiring professors who have been at the university for at least seven years, as both had been. According to Yukihisa Kitamura, an executive vice-president at Tohoku University, it was a "rare" measure taken while the two were under investigation for allegedly "dishonouring the university" by being involved with the website. In June, with the investigation still under way, the university granted the two emeritus status.

The unrest looks set to continue. During April and May last year, three key science-funding

agencies — the science and education ministry, the Science and Technology Agency and the New Energy and Industrial Technology Development Organization — all asked Tohoku University to evaluate accusations raised by Omura and three others. Citing confidentiality rules, a lawyer representing Inoue and Tohoku University declined to say what actions the university would take in response to the requests. Zhang has not responded to *Nature's* request for further comment.

Inoue has recognized some points raised by critics. In January last year, he published errata¹⁴ on two of the papers^{2,3} in dispute. In an accompanying explanation¹⁵, Inoue and his co-authors "apologize that these papers did not spell out all the experimental details as precise [*sic*] as we should have". But with the critics and the university at an apparent stalemate, the groups will probably be left to fight it out among themselves. In June 2010, Inoue sued Omura and others for defamation based on allegations posted on their website. In August, Omura's group countersued, claiming that Inoue was unjustly suing and trying to restrict their right to challenge research, and claiming defamation based on certain documents released by Inoue when initiating his lawsuit. The cases are pending.

Japan has no external body, akin to the US Office of Research Integrity, for investigating alleged scientific misconduct, despite calls for one from some quarters¹⁶. Only the minister of education has authority over a university president, and Kosaku Okada, a representative from the ministry's division familiar with the case, says that the ministry will not get involved. "We see the president as just another scientist and so we leave it up to the university to do any investigation," he says. But an independent investigation may be the only way to silence the critics. ■

agencies — the science and education ministry, the Science and Technology Agency and the New Energy and Industrial Technology Development Organization — all asked Tohoku University to evaluate accusations raised by Omura and three others. Citing confidentiality rules, a lawyer representing Inoue and Tohoku University declined to say what actions the university would take in response to the requests. Zhang has not responded to *Nature's* request for further comment.

Inoue has recognized some points raised by critics. In January last year, he published errata¹⁴ on two of the papers^{2,3} in dispute. In an accompanying explanation¹⁵, Inoue and his co-authors "apologize that these papers did not spell out all the experimental details as precise [*sic*] as we should have". But with the critics and the university at an apparent stalemate, the groups will probably be left to fight it out among themselves. In June 2010, Inoue sued Omura and others for defamation based on allegations posted on their website. In August, Omura's group countersued, claiming that Inoue was unjustly suing and trying to restrict their right to challenge research, and claiming defamation based on certain documents released by Inoue when initiating his lawsuit. The cases are pending.

Japan has no external body, akin to the US Office of Research Integrity, for investigating alleged scientific misconduct, despite calls for one from some quarters¹⁶. Only the minister of education has authority over a university president, and Kosaku Okada, a representative from the ministry's division familiar with the case, says that the ministry will not get involved. "We see the president as just another scientist and so we leave it up to the university to do any investigation," he says. But an independent investigation may be the only way to silence the critics. ■

1. *Mater. Trans.* **34**, 1234–1237 (1993).
2. *Mater. Trans.* **36**, 1184–1187 (1995).
3. *Mater. Trans.* **37**, 185–187 (1996).
4. *Mater. Trans.* **39**, 1001–1006 (1998).
5. *Scripta Mater.* **54**, 321–326 (2006).
6. *Nature Mater.* **8**, 887–891 (2009).
7. *Acta Mater.* **57**, 1290–1299 (2009).
8. *Acta Mater.* **54**, 2975–2982 (2006).
9. *J. Non-Cryst. Solids* **351**, 2519–2523 (2005).
10. *Mater. Trans.* **38**, 473–477 (1997).
11. *Mater. Trans.* **48**, 3190–3192 (2007).
12. *Mater. Trans.* **50**, 2502–2503 (2009).
13. *Mater. Trans.* **50**, 2504–2506 (2009).
14. *Mater. Trans.* **51**, 196a (2010).
15. *Mater. Trans.* **51**, 196b (2010).
16. *Nature* **439**, 514 (2006).



LATEST NEWS



US states striving to cull grey wolves could sidestep science
go.nature.com/4zmmic

MORE STORIES

- Connect the quantum dots for a full-colour image go.nature.com/oznkt
- Genetic engineering brings cloned crops closer go.nature.com/3tonwl
- Laser in reverse go.nature.com/sloyse
- Bone hormone affects male fertility go.nature.com/sqd47e

Q & A



Incoming director of developing-world science academy talks about brain drain
go.nature.com/ifhbbs

PHARMACEUTICALS

Genzyme deal set to alter biotech landscape

Maintaining innovative culture will be a key challenge for new owners of industry stalwart.

BY HEIDI LEDFORD

When the biotech company Genzyme first opened for business in 1981 in Cambridge, Massachusetts, it shared a building with a discount women's clothing store. "In the elevators you would see women looking for bargains, mixed in with these corporate types in suits, and technicians carrying mice in cages," recalls Harvey Lodish of Cambridge's Whitehead Institute, who was one of the company's founding scientists.

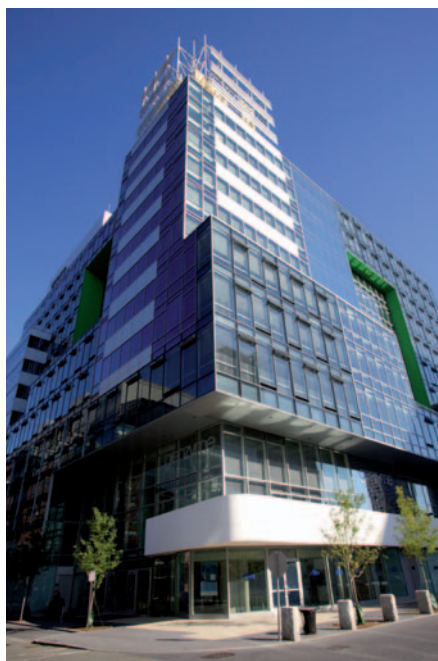
Thirty years on, with 10,000 employees, US\$4 billion in annual sales and a towering glass edifice for its head office, Genzyme is about to start a new life as a branch of the pharmaceutical giant sanofi-aventis. The \$20-billion deal, announced on 16 February, is the industry's biggest in two years. But some observers fear that the company's shift to the status of a branch plant will dilute the home-grown innovative culture for which Genzyme is known, and drain some of the dynamism from the biotech hub it has helped to create in the area around Boston.

Genzyme proved that designing therapies for rare diseases can be profitable. Early on, Genzyme scientists developed Ceredase, an enzyme obtained from human placentas to treat Gaucher's disease, an enzyme-deficiency disorder, and later followed it up with Cerezyme, a similar enzyme grown in mammalian cell cultures. It is Genzyme's expertise with therapeutic enzymes and experience with rare diseases that attracted sanofi.

The pattern has become a familiar one in recent years: a pharmaceutical giant acquires a smaller biotech company as a way of stocking up on new drug candidates (see 'Big buys'). The risk is that the culture clash between a larger, mature firm and its smaller, leaner acquisition can chase away the talent that the acquiring company was hoping to tap, says Josh Lerner, who studies entrepreneurship at Harvard Business School in Boston.

"Everyone's nervous" about the Genzyme deal, says Lodish, who no longer has a stake in the company. "It's a big unknown."

Genzyme is hardly a young start-up, but it is still more nimble than big pharma, says Alison Taunton-Rigby, chief executive of RiboNovix in Natick, Massachusetts, and a former senior vice-president at Genzyme. "No matter what



Genzyme: a cornerstone of Boston-area biotech.

big pharma does, it is different from biotech, and different people are drawn to these two different cultures," she says.

When Swiss pharmaceutical maker Roche fully acquired Genentech in 2009, for example, a mass exodus of talent ensued. Genentech, headquartered in South San Francisco, California, was renowned for its highly innovative culture, which attracted a steady flow of top-calibre scientists from academia. Despite Roche's assurances that its new acquisition would remain largely independent, many of Genentech's top executives left soon after the purchase was announced. "Since then, talented

people from Genentech have peppered the whole West Coast with biotech start-ups," says Neil Solomon of the Neil Michael Group in Great Neck, New York, an agency that recruits senior executives for life-sciences companies. Industry watchers say it is still too soon to gauge the impact on innovation at Genentech.

TROUBLE BREWING

It is a little early to see CVs coming from the top at Genzyme, Solomon says, but he has had conversations with executives there who are preparing their exits. This could portend trouble for sanofi, which is relying on Genzyme's expertise to strengthen its presence in protein-based therapies and the rare-diseases market.

Acquisitions by pharma do not always sound the death knell for innovation at biotechnology firms. Some say that Millennium Pharmaceuticals, another Cambridge biotechnology firm, may benefit from its acquisition by Japanese drug maker Takeda in 2008. Millennium had assembled a talented team of scientists, but its approach — to use genomics to fuel drug discovery — had foundered. In this case, the acquisition by Takeda could provide a turnaround by changing the focus or management of the company, says Gary Pisano of Harvard Business School, whose research includes the biotechnology industry.

What is clear is that the Genzyme acquisition heralds a change in the Cambridge biotech scene, says Taunton-Rigby. Genzyme and Biogen Idec, another stalwart biotech located only a few blocks away, have been the pillars of the local biotechnology community. Pressure from investors drove Biogen to seek a pharmaceutical suitor, and although none has yet come forward, the company has been cutting costs and jettisoning programmes. Just as Genzyme has become the rare-diseases division of sanofi, Biogen seems to be preparing itself to become the neurology division of a major pharmaceutical firm. These changes, says Taunton-Rigby, would sever Boston's connection to other aspects of drug development, such as marketing and production.

"The cornerstone of the business here has been these two fully independent companies," says Taunton-Rigby. "That's what made the culture here unique — when you only have research without the other capabilities, it's a very different scene." ■

BIG BUYS

The top five biotechnology acquisitions by pharma.

Biotech	Pharma	Price (US\$ billion)	Year
Genentech	Roche	46.8	2009
Genzyme	sanofi-aventis	20.1	2011
MedImmune	AstraZeneca	15.6	2007
Serono	Merck KGaA	13.3	2007
Millennium	Takeda	8.8	2008

K. MA/BLOOMBERG VIA GETTY IMAGES

SOURCE: COMPANIES

RESEARCH

Europe's innovation hub finally KICs off

Large networks of researchers and companies aim to develop and commercialize marketable products.

BY NATASHA GILBERT

In many ways, the European Institute of Innovation and Technology (EIT) is a classic European compromise. What started out as an effort to build an innovation powerhouse to rival the Massachusetts Institute of Technology (MIT) in Cambridge has been transformed by years of political wrangling into a gossamer web of academic and industry partnerships.

But some research policy-makers are pinning their hopes on the EIT as a cornerstone of the big push by the European Union (EU) to commercialize research and boost industry research and development (R&D) spending. Last week, after a long and painful gestation, the last piece of the EIT fell into place as its governing board gave the green light to a third and final Knowledge and Innovation Community (KIC).

This KIC aims to develop technologies for mitigating and adapting to climate change, and has a total budget of €360 million (US\$490 million) over four years. A vast — some would say unwieldy — network, it includes 20 core members and 68 affiliate partners, including French company EDF Energy and the City of Zurich in Switzerland. It follows previous KICs, agreed last December, on information and communication technologies, and clean energy. In all three cases, 25% of the funding comes from the EU budget and the rest from partner organizations (see 'Innovation partnership').

The EIT has been controversial ever since European Commission president José Manuel Barroso proposed the idea in 2005. Barroso imagined a centralized campus like MIT, but others argued that it would take many years for a single institution to build itself up from scratch to rival MIT. So the focus shifted to one of Europe's favourite fallbacks — a distributed network of partners.

Even this proved challenging to establish. The agreements for each of the three KICs took about a year to draft and approve. Mary Ritter, the climate KIC's chief executive and pro-rector for postgraduate and international affairs at Imperial College London, partly blames the "risk averse" nature of the European Commission, which allocates and audits EIT funding. But she also points out that the EIT has had to break new ground in establishing such large public-private networks.

The aim of the climate KIC, she says, "is to develop the most efficient means of producing and using fuel and other consumables, such as water, in a carbon neutral or negative way". The research will be carried out in five national centres across Europe, including Imperial in the United Kingdom and the Paris-Saclay R&D campus in France. The technologies developed here will then be tested and commercialized in one of the KIC's six regional innovation centres.

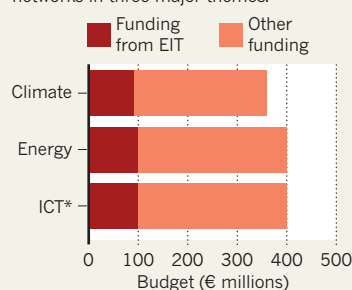
Some members of the European Parliament say that they will push for the EIT to receive a greater share of the EU budget in future, in part to help deliver the 'Innovation Union' agenda being promoted by Europe's research commissioner Máire Geoghegan-Quinn to translate more of Europe's basic research into marketable products.

Katrien Maes, chief policy officer at the League of European Research Universities, based in Leuven, Belgium, applauds the EIT's goal of boosting innovation. But she complains that the EU has dictated the EIT's research agenda from the top down, rather than giving it the flexibility to set its own targets — something it needs to be truly innovative.

The future direction of the EIT will be discussed in April at a conference in Budapest, where the institute has its headquarters. There is a lot riding on the success of the first three KICs: the EIT will fund more only if these fore-runners bear fruit. ■

INNOVATION PARTNERSHIP

The European Institute of Innovation and Technology (EIT) is supporting research networks in three major themes.



*Information and communication technology

SOURCE: EIT

NATURE.COM
Read more at Nature
Europe:
go.nature.com/ptxmzt



B. CANADAY/HARVARD UNIV.

Q&A Joe Aldy

Climate strife and political life

Joe Aldy was special assistant to President Barack Obama for energy and environment, and represented the White House during climate negotiations in Copenhagen in 2009 and in Cancún, Mexico, in 2010. He stepped down last December and now teaches public policy at Harvard University's John F. Kennedy School of Government in Cambridge, Massachusetts. As a climate debate shapes up in the House of Representatives, he offers Nature some perspective on the legislative battle ahead.

What effect is the Republican majority in the House of Representatives having on climate policy in Obama's administration?

House Republicans want to strip the Environmental Protection Agency (EPA) of its climate regulatory authority and cut clean-energy and climate funding. Some of this may be theatre, but it could drive hard negotiating positions over the budget and even shut down the government if no agreement can be reached. International climate funding could be a casualty, with a compromise delivering less than the administration's budget request and less, even, than what some conservative groups support.

How far can the administration go with the EPA regulations?

The EPA can make some progress in reducing

emissions through the Clean Air Act, but the act is not well designed for tackling climate change. The president has established ambitious long-term climate goals, and if the country is going to meet those in a cost-effective manner, I think we need legislation.

How should the administration move forward in the face of Republican opposition?

I'm cautiously optimistic that this Congress can move some form of climate-relevant legislation. If the administration can secure a meaningful clean-energy standard in the power sector, then I think it makes sense to exchange it for EPA regulatory authority. A good clean-energy standard could potentially remove the need for additional regulations.

Could the White House have pushed harder to pass climate legislation last year?

Part of this is a function of where Congress is and what Congress wants to do. In the previous Congress, the House moved forward with energy and climate first. If the Senate had done so, I think the president could have pushed it over the finish line and secured a bill. It probably would have meant that the clock would have run out for health care.

What message do you take from the climate bill's failure?

It is important that legislation is kept simple and transparent. A very complicated bill could undermine what the administration is trying to achieve in terms of driving investment in new energy technologies.

What were your thoughts when a deal struck between the United States and several key nations at the 2009 Copenhagen climate talks did not win broader support?

There is no doubt that the UN plenary was a painful process to watch. It was frustrating to know that an agreement was reached between more than 25 heads of state representing every major economy and every group in the UN negotiating process, and yet the agreement was still being derided as illegitimate and grossly inadequate. And although there was more that the administration would have liked to secure in the deal, it certainly beat having nothing. This is one reason why the meeting in Cancún was important, because a deal was struck there that incorporates all of the major elements that the world leaders agreed to in Copenhagen.

What are those elements?

First, that developed and developing countries are standing side by side to reduce their emissions. Second, that there is serious support for those who are the poorest and the most vulnerable to a changing climate. And third, there is an agreement that will establish how the world is going to review, analyse and consult on all of the major mitigation actions that the major economies are taking. This will make it easier to figure out what policies work, which I think is very, very important.

Kyoto Protocol commitments by industrialized nations will run out next year. Does Kyoto have a future?

Kyoto doesn't necessarily disappear, but the idea of splitting the world into developed and developing economies on this issue seems destined to disappear. More than half of emissions are coming from the developing world. If we don't adjust to that, then everything we are negotiating and all of the actions that will be driven by these negotiations will be inadequate to the challenge of climate change. ■

INTERVIEW BY JEFF TOLLEFSON

In the past five years animal-rights activists have perpetrated a string of violent attacks. In February 2008, the husband of a breast-cancer biologist in Santa Cruz, California, was physically assaulted at the front door of their home. In the same month, the biomedical research institute at Hasselt University in Diepenbeek, Belgium, was set on fire. In the summer of 2009, activists desecrated graves belonging to the family of Daniel Vasella, then chief executive of the pharmaceutical company Novartis, based in Basel, Switzerland, and torched his holiday home.

A poll of nearly 1,000 biomedical scientists, conducted by *Nature*, reveals the widespread impact of animal-rights activism. Extreme attacks are rare, and there does not seem to have been any increase in the rate of their incidence in the past few years, but almost one-quarter of respondents said that they or someone they know has been affected negatively by activism.

More than 90% of respondents agreed that the use of animals in research is essential, but the poll also highlights mixed feelings on the issue. Nearly 16% of those conducting animal research said that they have had misgivings about it, and although researchers overwhelmingly feel free to discuss these concerns with colleagues, many seem less at ease with doing so in public. More than 70% said that the polarized nature of the debate makes it difficult to voice a nuanced opinion on the subject, and little more than one-quarter said that their institutions offer training and assistance in communicating broadly about the importance of animal research (see 'Assessing the threats').

CRACKING DOWN

During the past decade, both the United States and the United Kingdom have enacted tough laws in response to violent tactics from activists. In 2005, the United Kingdom created the Serious and Organised Crime and Police Act, allowing stiff sentences to be imposed on those who intimidate companies and individuals that contract with animal-testing labs. Activists have since been found guilty of blackmail for terrorizing individuals and companies with financial ties to Huntingdon Life Sciences, a contract animal-testing company in Cambridgeshire, UK (see page 454). In the United States, the 2008 Animal Enterprise Terrorism Act was brought in to combat property damage and threats that produce a 'reasonable fear' of death or injury for researchers or their relatives, although its enforcement has been challenged in the courts.

These laws do not seem to have driven down the rate of violence. The Foundation for Biomedical Research in Washington DC, which is in favour of animal research, and the anti-animal-research magazine *Bite Back*, based in West Palm Beach, Florida, collect accounts of activism incidents from media reports and activist websites, respectively. Although not comprehensive, their data suggest that the worldwide incident rate has been stable for five years or more, with some regional variation. Activity in Britain seems to have dropped since the anti-Huntingdon campaign cooled. Protests have also been scaled back at the Biomedical Sciences Building at the University of Oxford, which opened in 2008 and houses research animals including primates.

Nearly one-quarter of biologists say they have been affected by animal activists. A *Nature* poll exposes the

BATTLE SCARS

BY DANIEL CRESSEY

Although *Nature's* survey was not designed to measure the incidence of activism, it suggests a similar picture: 45% of respondents said they had not perceived an increase in activist activity in the past five years, with some regional differences. US scientists were more likely to say that activism had increased, whereas many UK scientists reported a perceived decrease. Sally Rockey, deputy director for extramural research at the US National Institutes of Health in Bethesda, Maryland, says that the responses probably reflect the publicity drawn by high-profile incidents, not real increases. "There have been some life-threatening situations, arson and bomb threats for example. One of the things we've seen is some investigators have been targeted at their homes," says Rockey.

Animal researchers who said that they

or someone they knew had been affected by activism wrote about incidents ranging from anonymous threats and protests outside laboratories to vandalism, 'liberation' of animals, physical attacks by masked activists and bombs both real and simulated. "Home damaged, young children terrorized, death threat, etc," reports one genomics researcher matter-of-factly.

A small number, about 15% (26 respondents), who had been negatively affected by activism said that they had changed the direction or practice of their research as a result. After encountering violent protests, one US academic was "much less willing to conduct any studies on non-human primates, despite their absolute critical relevance for neuro-protection research".

PRIMAL CONCERNS

Only 38 scientists working with non-human primates responded to the survey, but they were the group of respondents most likely to strongly agree that activism is a problem. Frankie Trull, president of the Foundation for Biomedical Research, says that in her experience, primate researchers are targeted more than those in any other type of animal work.

Although more primate research is conducted in the United States, the ability to work with primates has been challenged in Europe. In 2009, the European Union considered legislation that would have restricted work on non-human primates to research investigating "life-threatening or debilitating" conditions. It took a concerted campaign by researchers to amend it to allow for basic research in addition to applied work.

Hannah Buchanan-Smith, an animal-welfare researcher at the University of Stirling, UK, says, "Primate laboratory researchers are finding it harder to justify their research to the public." Buchanan-Smith refuses to do any animal research that causes pain, suffering, distress or lasting harm, and says that basic research on primates presents a particular ethical challenge.

She argues for alternatives to animal testing. "Replacement is the ultimate goal and we are moving in that direction with certain groups of animals," she says. "I very much hope in my lifetime that will be achieved in primate research."

Stefan Treue, head of the German Primate Centre in Göttingen, views primate research in a different light.

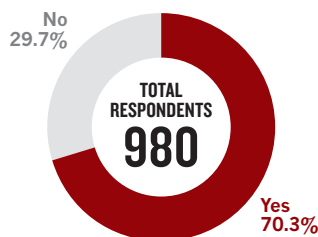
He says that after lay-people have visited his laboratory and seen how work

**ANIMAL
RESEARCH**
nature.com/animalresearch

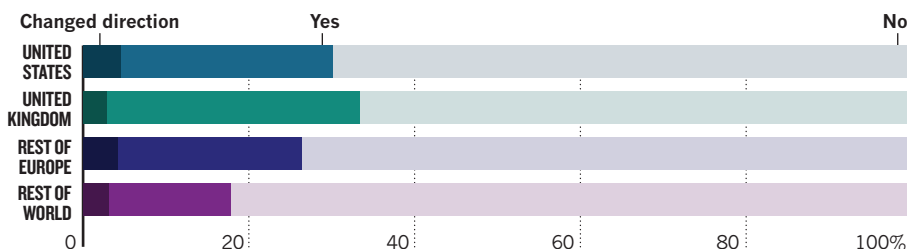
Assessing the threats

Biomedical scientists who responded to a survey mostly support animal experimentation, but attitudes towards animal-rights activists differ around the world.

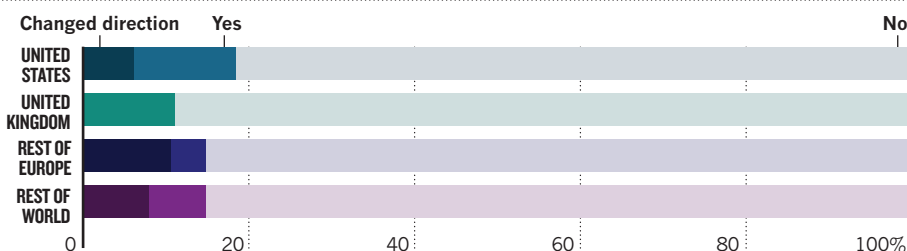
Q Do you conduct experiments on animals?



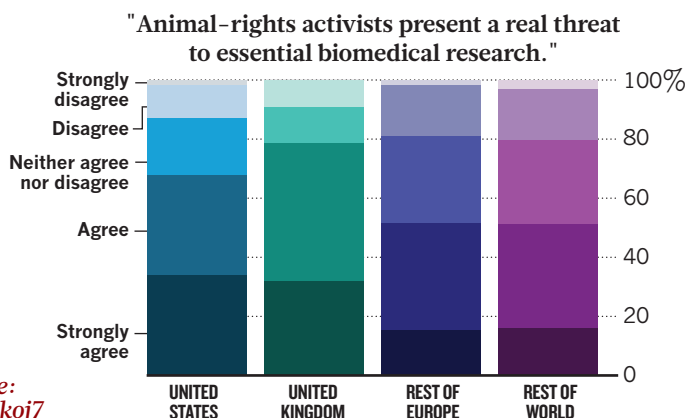
Q During your career, have you or anyone you know been affected negatively by animal-rights activists and have you changed research direction as a result?



Q Have you ever had misgivings about the role of animals in your research, and did you change research direction as a result?

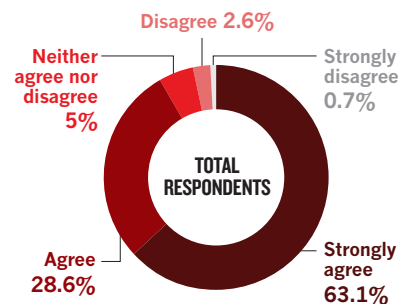


Q To what extent do you agree or disagree with the following statements?



For full results, see: go.nature.com/o6koj7

“Animal research is essential to the advancement of biomedical science.”



is conducted and why, “something like 98% understand and accept that this is a small but important and irreplaceable part of biomedical science that is conducted to the highest ethical standards”.

Treue rejects an ethical distinction between basic and applied research. “It’s not a logical argument to say, ‘I accept applied research but I don’t want the underlying basic research’, because you can’t have one without the other. I have to admit that partly the science community is to blame for not explaining that more clearly and more frequently in public,” he says.

PUBLIC CONSULTATION

Some results from the survey suggest that communication with the public might be improving. Fifty-five per cent of animal researchers said that their institutions encourage communication with the general public about their work, and only 7% said that this is actively discouraged. In a

poll run by *Nature* on this subject in 2006, only 29% of researchers said that they were encouraged to discuss their work, and 11% had been discouraged (see *Nature* 444, 808–810; 2006).

This is good news, says Rockey, but there is much to be done. More than half of the researchers who said they are encouraged to discuss their work indicated that their institutions offered no support or training on how to do so. “It’s important for institutions to have outreach programmes which engage the public in explaining the importance of the research,” says Rockey.

It can be challenging to explain the type of nuanced positions on animal research that the poll revealed: 33% of respondents had “ethical concerns” about the role of animals in their current work. Researchers wrote about their preoccupations with reducing pain, minimizing the numbers of animals used and showing respect for their subjects. Some 16% reported “misgivings”

about work they have done, and half of these (54 researchers) said that they changed their research or practices as a result, suggesting that personal reflection may be more effective than activism in changing behaviour. “I consider these issues virtually daily,” wrote a US neuroscientist. “The day I stop considering these issues is the day I quit. I know few scientists who don’t feel similarly.”

Trull welcomes scientists thinking deeply about the issues involved in working with animals, and is glad that 93% of researchers said they feel free to discuss concerns about ethics with colleagues. “There are a lot of those discussions and debates that go on in the research community. It’s a privilege to use these animal models,” she says. “Scientists need to view it in this way and I think they do.” ■ SEE EDITORIAL P.435

Daniel Cressey writes for *Nature* from London. Survey work was aided by **Laura Harper**.

THE RADICAL

Researcher by day and activist by night, Joseph Harris was leading an untenable double life that eventually landed him in prison.

BY SHANTA BARLEY

Most British PhD students head straight to the pub after graduating. Joseph Harris had something else planned. On a freezing Friday night in December 2005, he packed a rucksack with a pair of bolt cutters, a hammer, a flashlight, latex gloves and a can of black spray paint. He drove to a deserted industrial estate just outside Nottingham, UK, and broke into a compound owned by a company that hires out refrigeration units. He slashed the tyres of a van and glued down its windscreen wipers. He smashed his way into an outhouse, tipped cans of lubricant and coolant onto the floor and cut through the electrical cables that powered the office's air-conditioning unit. On his way out he sprayed, "Now you pay 4 your crimes" on a wall.

The following Monday, Harris returned to work in Sue Watson's oncology lab at Queen's Medical Centre in Nottingham, where he was already committed to an irreconcilable double life. By day, he studied potential treatments for gastrointestinal cancer — work that invariably required the use of animal models. By night, he crusaded against such animal research, sabotaging companies with links to it. Within a month, Harris would be caught vandalizing another company. Ultimately, he would become the first person in the United Kingdom to be convicted under a law intended to crack down on activist extremism.

The judge overseeing the case lamented a career destroyed and a scientist lost. "It may well be that your future inability to continue your research into gastrointestinal cancer will be a great loss to those who suffer that disease," said Judge Ian Alexander when he sentenced Harris to three years in prison. Harris was released within a year, and has now agreed to speak to *Nature* about what drove him to commit his crimes.

RUMMAGING FOR ROLY-POLIES

Most of Harris's childhood in southern England's New Forest was spent outdoors. He and his younger brother, Thomas, passed the hours rummaging under rocks for roly-polies — woodlice that roll into a ball when threatened. Joseph Harris showed an early affinity for science. At the age of ten he started compiling a list of every animal in the world. "I didn't get very far," he says. His convictions were already firm, however. In the same year, his teacher asked him to write about what he wanted to be when he grew up. Instead, he wrote about what he didn't want to be: a scientist who did tests on animals.

He excelled in biology and chemistry at school, and in 1998 began a degree in molecular biology at the University of Nottingham. While there, Harris joined the local Hunt Saboteurs Association group in fox-hunt sabotage, or 'sabbing'. Shadowing hunts from Land Rovers, he and his friends sprayed citronella to mask the scent of foxes and break pursuit. His actions were legal as long as he didn't injure anyone or trespass on private property.

In 2001, Harris was interviewed for a PhD on oesophageal cancer at Queen's Medical Centre. "I made it clear that I wouldn't do animal testing during the interview, even though I thought it would scupper my chances," says Harris. Instead, he got the job, with a grant from the UK Medical Research Council. He soon found himself in Watson's lab studying a class of drugs known as proton pump inhibitors. These drugs are commonly prescribed for gastro-oesophageal reflux disease, but tests on rats in the

1980s suggested that they might also boost the risk of developing a type of cancer known as oesophageal adenocarcinoma (N. Havu *Digestion* 35, 42–45; 1986). According to one hypothesis, they do this by raising the amount of the growth hormone gastrin in the bloodstream. Harris and his colleagues suggested how, by showing that gastrin can suppress a cell-death pathway in oesophageal cells, allowing them to survive and possibly become cancerous (J. C. Harris *et al. Cancer Res.* 64, 1915–1919; 2004).

"I wasn't doing any animal testing," says Harris, "but I still felt that I was making important contributions towards the field."

PETTY INSOLENCE

As his involvement in research gained momentum, Harris says that his participation in activism waned, save for a single act of petty insolence. One day in a nearby lab, he spied a pair of axolotls — amphibians studied for their ability to regrow entire limbs. With no one looking, he decanted them into a jar and took them home. "I just felt so sorry for them," he says. But as his doctoral studies neared an end, he was forced to confront the conflict between his career and his convictions. Watson urged Harris to take on the project of a postdoc who had quit the lab. This meant an early promotion for Harris — still technically a student — but the project involved working with mice. "When I took over, she asked me, gently and then increasingly insistently, to continue the animal testing, even though she knew that I was utterly against it," he says. The pressure was real, says colleague Jacqueline Dickson, who also did her PhD under Watson's supervision. "Sue is very used to getting her own way," says Dickson. Watson declined to comment for this story.

Eventually, Harris acquiesced. He was studying whether a drug designed to prevent gastrin from binding to cell receptors would slow tumour growth in mice. Harris says that he reluctantly agreed to analyse biopsies. "I didn't inject a single drug, or see a single animal die, but I still felt complicit in what was happening," he says. Worst of all, Harris was concerned that the animals were dying for nothing. Phase III clinical trials had shown that the drug, known as Insegia or G17DT, did not improve survival rates in patients with pancreatic cancer, and Harris reasoned that it probably wouldn't work for gastrointestinal tumours either. He was appalled, he says, that his own research into gastrin had been used to justify the drug's continued development.

Tortured by his conscience, Harris felt the need to do something. Huntingdon Life Sciences (HLS), the largest animal-testing company in Europe, located in Cambridgeshire, UK, seemed like an appropriate target. He never forgot the highly publicized release of video footage secretly shot at HLS in 1997, showing several severe breaches of animal-research protocols, including beagles being beaten. The films had already riled the broader British activist community into a state of organized fury. Activists firebombed three cars owned by HLS staff and sent the managing director mouse traps allegedly tainted with HIV — he was later attacked by three men armed with pickaxe handles. The organization Stop Huntingdon Animal Cruelty (SHAC), which formed in 1999 and is based in London, targeted HLS investors relentlessly with forms of intimidation such as hoax bombs and abusive e-mails. One by one, financial backers of the company

K. HARVEY/NATURE

**ANIMAL
RESEARCH**
nature.com/animalresearch



pulled out, fearful that their staff would be hurt. In early 2001, the Royal Bank of Scotland cancelled a £22-million (US\$47 million in 2001) loan, forcing HLS to near bankruptcy.

But HLS and the BioIndustry Association — an organization based in London that promotes UK biotechnology — fought back, lobbying the government for tougher laws against activists. The London-based Association of the British Pharmaceutical Industry delivered an ultimatum: if the government did not do something to curtail activist attacks, Britain's drug industry would boycott those banks and investment companies that had caved under the pressure of SHAC. "All of a sudden, the government was confronted with the real threat of pharmaceutical companies pulling out of future research-and-development investment in Britain," says Andrew Upton, who studied the animal-rights movement for his PhD at Liverpool John Moores University, UK. The UK government passed the Serious and Organised Crime and Police Act in 2005. Section 145 of the law explicitly prohibits "interference with contractual relationships so as to harm animal research organisation", and imposes tough sentences of up to five years for perpetrators of these crimes.

CRIME AND PUNISHMENT

The threat of punishment wasn't enough to dissuade Harris — but the armed patrols and barbed-wire perimeter now installed at HLS did convince him to go after a softer target. Like several activists before him, Harris went after companies that had contracts with HLS (and much less security). The SHAC website had identified several such companies, and for his first foray, Harris chose the Nottingham branch of York Refrigeration, which provided fridge units to HLS. The fact that he had graduated with a PhD that very day was "just a coincidence", says Harris.

Sensing no repercussions from his acts of anarchy, he resumed his crusade one month later. On 15 January 2006, Harris targeted Atlas Material Testing Solutions, a company based in Bicester, UK, that provided HLS with equipment. He sliced through power cables, squirted superglue into locks and flooded the office. "The sight of a hosepipe propped up against the front door proved too much to resist," says Harris. He pushed the nozzle through the letter box, twisted the tap and destroyed a shipment of computers still in their Royal Mail boxes. Lastly, he spray-painted, "This company kills puppies" onto a wall.

Later that evening, he broke into Bullimore Plant Hire in Northampton, UK, which was helping HLS to build an extension. He was snipping air valves off the tyres of bulldozers when he heard a distant shout. Soon after, ten police cars arrived on the scene, responding to what officers thought was an attempted burglary at an adjacent electronics warehouse. Harris ran and the police gave chase. "I knew it was all over, but I didn't want to just give myself up," says Harris. Defiant to the last, he buried his hands beneath him when he was slammed to the ground to avoid being cuffed.

It took several months of police work to reveal that Harris was not, in fact, a burglar. DNA and minute droplets of copper on the blades of his bolt cutters eventually tied him to the crimes at the three companies, which had caused damages in excess of £28,000 (US\$49,000 at the time). Harris insists that these are the only serious crimes he has committed. After his arrest, the police raided his lab and took away his computer. He was subsequently fired. His actions shocked many of the staff at Queen's Medical Centre, but not Dickson. "Joe's a 'straight-edged vegan' who makes no secret of his opposition to animal testing, doesn't drink alcohol or coffee, and donates most of his money to animal-welfare charities, so although I was surprised it kind of all made sense," she says.

Harris was sentenced in September 2006 and spent his first month in prison in a maximum-security facility near Milton Keynes, alongside,

"He is the only example on record of an animal-rights activist actively contributing to the very thing he was trying to wipe out."

➔ NATURE.COM

For a podcast discussion with Joseph Harris, see: go.nature.com/9ajhgc

he says, murderers, rapists and drug dealers. Harris dead-pans that he learnt two valuable lessons in prison: how to stick a poster to a wall using toothpaste, and how to take a compliment. "A guy came up to me in the shower and told me I had J-Lo's bottom," says Harris. "I thanked him and walked away." He shaved off his curly chestnut hair to look tougher in front of the other inmates, but his concerns about being bullied were unfounded. If anything, he was, it seems, universally liked by his companions. When his image appeared on the news one evening his fellow inmates cheered, he says. They also expressed bafflement over his three-year sentence. "Meeting a member of the National Front who got the same sentence as me for stabbing a black man in the throat, killing him, made me wonder too," says Harris. He later appealed his sentence on the grounds that Judge Alexander was a fox hunter and Harris was a first-time offender. He was released in August 2007 after 11 months in prison.

PALPABLE FEAR

David Jentsch, a neurobiologist at the University of California, Los Angeles, who uses vervet monkeys to study the effect of drug addiction on the brain, doesn't think that the punishments imposed on activists are too severe. In March 2009 his car was firebombed, and in November last year he received a parcel containing HIV-tainted razors — with an accompanying threat that his throat would be slashed. Whether psychological or physical in nature, the violence is terrifying, he says. "You feel fear. Real, palpable fear." Harris describes the threats that Jentsch has received as horrendous, but justified. "Vervet monkeys are very intelligent creatures that, like humans, feel psychological fear, so each time David goes into the room they're probably terrified too," he says.

Harris represents a unique brand of animal-rights activist, according to Upton. "He is the only example on record of an animal-rights activist actively contributing to the very thing he was trying to wipe out," says Upton. Harris also operated alone and had a 9-to-5 job. Most animal-rights activists who have broken the law since 2005, says Upton, have been full-time protesters embedded in campaigns. "They were 'career' activists, for whom convictions were an occupational hazard," he says. Harris's brother Thomas fits this profile better. He dropped out of university after a year, and in 2007 took over running SHAC after several of its leaders were imprisoned. He is currently serving a five-year sentence for intimidating companies linked to HLS. Joseph maintains that he never spoke to his brother about his own escapades and that he acted independently of SHAC. He says that he found the members of the organization 'cliquey'.

In some ways, however, Joseph Harris does fit the activist profile: white, middle class and well educated. James Jasper, a sociologist at the City University of New York, studies animal-rights protesters and says that people in this category are more likely to become animal-rights activists because they do not need to fight for their own basic rights, and are therefore at liberty to extend rights to less privileged groups.

"I didn't really think of myself as an animal-rights activist," Harris says now. "I was just doing what I thought was right." Harris has tried to resume a normal life since being released on parole. Initially, he could only find a job driving forklift trucks in Newbury, UK. After eighteen months, however, he persuaded a well-known conservation organization to take him on as an unpaid intern. He asked that they not be named for fear that he might be forced out.

Today, Harris speaks about his past double-life as though it happened a long time ago. He knows he can never return to biomedical research and regrets that he may never be able to do a PhD in conservation. "It's hard to imagine that any professor would take on someone with my past," he says. But he is also unrepentant. Although he was caught, he views his crimes as a success. All three of the companies that he sabotaged pulled out of their contracts with HLS to prevent further attacks. "You can't get much more effective than that," he says. ■ [SEE EDITORIAL P.435](#)

Shanta Barley is a freelance writer based in London.

COMMENT

GEOSCIENCE A lesson for the future in a history of how life shaped Earth **p.460**



HISTORY How novels elevated sensibility above mechanical views **p.462**

CHEMISTRY Preserving the scent of endangered flowering plants **p.464**

PSYCHOLOGY On the definition, treatment and prediction of stuttering **p.465**

HARDOASH PRODUCTIONS



Scenes from the film *Monkeys, Rats and Me* show how animal research is done in Oxford, UK.

**ANIMAL
RESEARCH**
nature.com/animalresearch

Animal testing: TV or not TV?

Two views on whether scientists who believe that animal experimentation is necessary should become public advocates, or work quietly behind the scenes.

POINT

Somebody has to speak out

Tipu Aziz and John Stein, *John Radcliffe Hospital, Oxford, UK.*

It was not without trepidation that, in 2006, we appeared on prime-time national television to talk about our work on macaque monkey models of Parkinson's disease. After all, there have been severe repercussions for researchers in the United Kingdom. In 1985, Molotov cocktails were thrown at the home of Nobel prizewinner John Vane, then director of research at the Wellcome Foundation. In 2004, activists exhumed the body of Gladys Hammond, just because she was the late mother-in-law of a breeder of guinea pigs for animal experiments (this at least provoked public outrage). Colin Blakemore, a ►

COUNTERPOINT

Avoid the camera

Ranga Yogeshwar, *science TV presenter near Cologne, Germany.*

Television is how most people get their news these days: it brings remote lands closer, and provides emotionally compelling close-ups of individuals affected by world events. But it is a totally unsuitable platform for delivering complicated information or detailed discussion. My advice for scientists involved in the ethically complex field of animal research? Stay as far away from the camera as possible.

I say this as a science TV presenter and former scientist. I was once a particle physicist and am now on the board of a couple of Max Planck institutes, and I have hosted science television in Germany since ►

POINT: SOMEBODY HAS TO SPEAK OUT ► neurobiologist at the University of Oxford and former head of the UK Medical Research Council, is one of the few scientists who has been consistently brave enough to defend his animal research in public. The result for him has been assaults by masked terrorists, bombs sent to his children and letters laced with razor blades: in short, two decades of attacks and abuse.

Activist intimidation and violence has understandably led many scientists to conclude that it is not safe to speak in public about their work. This silence has opened the way for anti-vivisectionists to misrepresent animal experiments as useless, cruel and outdated. In one debate in which we were involved, the audience were told that no animals were used in the discovery of insulin or of penicillin. This is not true: insulin was discovered with the help of experimental dogs, and the first preparations of penicillin were tested on mice. Such bending of the facts is easy if there is nobody to present the other side.

The one-sidedness of the debate has had some disastrous effects. In 1997, the University of Cambridge, UK, announced amid great fanfare a plan for a new primate-research centre, supported by £24 million (US\$39 million) from the Wellcome Trust. But constant harassment of staff and members of the public by animal activists led to loss of local support, and in 2004 the university abandoned the project. Remarkably, not a single scientist who might have made use of the facility dared to defend the need for it in public.

We have been able to show that it doesn't have to be this way. In 2004, construction began on a large animal-research facility in Oxford that would also house primates. Soon after, extremists set fire to the contractor's offices, and construction was stopped for more than a year. Our research towards better Parkinson's disease treatments was in grave jeopardy. For some time we had been urging the University of Oxford's central management to be more proactive and explain to the public why animal experiments are so important — without success. So, after anxious debate with our families, despite the risks, we decided to speak up. Someone needed to do something. Otherwise we, too, would lose the battle.

INSULTS AND DEATH THREATS

We had the advantage that we could tell a fairly simple story about how our primate research was feeding directly through to improvements in neurosurgery, and we had patients who were so grateful that they too agreed to speak up. We thought this might help to counteract the misinformation being put out by SPEAK (an animal-rights group originally founded as Stop Primate Experiments at Cambridge (SPEAC), which had moved its attentions to Oxford).

We gave an interview to the *Oxford Times* in 2004 explaining that the new building was needed to provide better conditions for experimental animals, which are the basis for virtually all medical advances and ensure the safety of any new drugs. This earned us a storm of insults, death threats by e-mail and letter, and, sadly, strong criticism from many of our colleagues for 'putting their lives at risk'. *Nature* picked up what we were trying to do, however, and wrote a piece on the need for the facilities and the need for educating the public to appreciate their value (see *Nature* 430, 597; 2004).

We two scientists alone did not make much of an impact until a 16-year-old student, Laurie Pycroft, came upon a SPEAK march in Oxford. He bravely joined in with a banner supporting animal testing, until he was summarily ejected. He then set up a website, using his slogan 'Pro-Test', and rapidly found himself supported by hundreds of students from Oxford and elsewhere. We were honoured to be invited to join them. Together, we decided to organize a Pro-Test demonstration to bring the debate about animal research to the attention of the public.

Pro-Test led the first ever UK march in support of the benefits of



John Stein and Tipu Aziz at the first Pro-Test rally, in Oxford, UK, in 2006.

animal research, which we were proud to lead with our local Member of Parliament, Evan Harris. We were joined by more than 800 supporters, which was, as *The Times* reported, five times more than SPEAK could muster. We were pleased by the considerable media attention that this gained, and contacted the BBC with an idea to make a documentary about our and others' animal research in Oxford. With some difficulty we managed to overcome the objections of both the hospital and the university authorities against allowing filming in our animal laboratories. The end product, *Monkeys, Rats and Me* with journalist Adam Wishart, aired in November 2006. This led to a proper national discussion. A debate was featured on the renowned BBC current-affairs programme *Newsnight*. And the subject of animal experimentation is now often in the 'general studies' curriculum of students hoping to move on to university.

By the end of 2006, former prime minister Tony Blair had committed the Labour Party to the need for animal research, and work on the Oxford facility began again under utmost security. During this period, injunctions were issued against known activists, banning them from approaching the construction site and the residences of scientists involved in animal research, including us. The police were impressively efficient and we felt well protected. By November 2008, the laboratory was completed. The state-of-the-art facilities are admired by almost everyone except for the small bunch of anti-vivisectionist protestors who still demonstrate outside every Thursday.

Unfortunately, the immense security and very high cost of macaques, which, in the United Kingdom, have to be purpose-bred if needed for research, have led to as much as a fifty-fold increase in the cost of carrying out primate research in the country over the past 20 years. This, more than activists' direct threats, is driving research abroad, often to countries where animal welfare is less scrupulously protected. If activists really care about animal well-being they should actually encourage the work to take place in the United Kingdom.

In this electronic age, remaining silent is not an option for scientists engaged in animal research; anyone can use Google to look up what a researcher does. The sensible thing is to be proactive and prepared to defend our work. The public needs to hear all sides of the story. Every scientist must make his or her own moral judgements on this issue, but we would make the same decision again. ■

Tipu Aziz is in the Department of Neurosurgery, John Radcliffe Hospital, Oxford OX3 9DU, UK. **John Stein** is in the Department of Physiology, Anatomy and Genetics, Oxford OX1 3PT, UK. e-mail: john.stein@dpag.ox.ac.uk

COUNTERPOINT: AVOID THE CAMERA ► the 1980s. One programme that I write and host, *Quarks & Co*, gets some 1.5 million viewers a week, and another, *Die große Show der Naturwunder* (The Great Natural Wonder Show), is aired four times a year and gets one of the highest viewing figures of any show in Germany: about 5 million to 8 million people.

My goal is not to court ratings, but to present content honestly. I try to fight against the current trends of over-simplifying the news, over-playing the implications of research or chasing human-interest stories above all else. But I know that some stories simply don't work well on television. I can remember only one time that my show has tackled animal research, and even then the point was not to show its necessity, but to highlight work on alternatives. In that way we could portray scientists as 'the good guys'.

The problem is that on television, emotionally-charged pictures rule. Scenes of seal cubs clubbed to death precipitated a worldwide campaign against furs. It was gruesome pictures coming out of Vietnam that led to the end of that war. By the same token, film sequences of animal experimentation trigger intense reactions in viewers. Scientists are charged with cruelty and the accusation cannot be countered with arguments: the screech of a monkey or whine of a dog drowns out every possible justification. Researchers I know have told me that they themselves don't want to watch scenes of animal experimentation. The typical reaction is disgust — in my opinion, for good reason.

The ethics of animal experimentation are not simple. Although it is clearly a part of advancing medical science, there are times when its use is debatable. In the pharmaceutical industry, some new drugs are developed because they are expected to provide a financial payoff, even though effective drugs for the condition in question already exist. We might ask whether it is really justified to kill animals for that. I once worked at a nuclear reactor where cancer research was being done. I had an argument with one scientist because I thought that he was working in such a sloppy way that much of his data, produced at the expense of dozens of rats, meant nothing. Doing an animal experiment needs as much care as landing a jumbo jet: it requires not just a system of strict rules and regulations, but also personal dedication to best behaviour on behalf of everyone involved.

Even in the academic world, there is no consensus about what counts as good, necessary animal research. This kind of moral complexity does not come across well on television. Viewers can and do seek refuge by turning to another channel if a programme is boring or demanding. The length of stay on a given channel has dropped spectacularly

over the past few decades, in part because of the introduction of dozens more channels — for *Quarks & Co*, a 45-minute show, viewers watch on average 23 minutes. In response, the mass media tries to be clear-cut and simple. It draws a line between good and evil, culprit and victim, winner and loser.

An audience presented with multiple opinions and viewpoints tends not to weigh the credentials and expertise of the speakers, but to decide on instinct who is more emotionally credible. Even if a television programme is convinced to not air graphic images of animals, an actress weeping about the fate of a puppy will carry more weight than a dry scientist with a logical defence. And, to top it all, scientists are often poor communicators.

In cases where scientists fear that their funding will be cut or their lab closed down, they should make their case to politicians, or to the public, through the less-emotional medium of print reporting. In cases where scientists face fanatical activists, that threat will always remain, whether

the researchers speak up about their work with animals or not. No amount of argument will change such views.

The general public, on the other hand, are not a threat. They know that animal experiments exist, just as they know how meat is produced and how battery-farm chickens are treated. But most are happy to turn a blind eye to these things and accept the benefits, from pharmaceuticals to steak and eggs. This is normal, and scientists can use it to their advantage.

Many things that are good for society as a whole would not survive a plebiscite on how public funds should be spent. Against the backdrop of worldwide budget cuts, an open discussion about the cost-benefit ratio of high-energy physics (or opera) might not be good for the field.

We should not encourage television to draw attention to debates that can't be won, or to topics that do better left entirely out of the spotlight. Anyone who makes the mistake of trying to present the complex story of animal experimentation through the colourful lens of television will end up in darkness. ■

Ranga Yogeshwar is a science journalist based near Cologne, Germany. He presents several television programmes in Germany and writes popular science books, including *Ach, so!* (2010). e-mail: info@yogeshwar.de

DIE GROSSE SHOW DER NATURWUNDER/SWR



Only some animal research comes across well on television: Ranga Yogeshwar (right) with a raccoon on *Die große Show der Naturwunder*.



J. READER/SPL

Earth was transformed by the emergence of oxygen-producing photosynthetic organisms — such as the cyanobacteria found in these ancient stromatolites.

EARTH SYSTEMS

Shaped by life

Wolfgang Lucht sees a lesson for humanity's future in the long co-evolution of our planet and its inhabitants.

Earth has life written all over it: in its sediments and soils, its atmosphere and its oceans. Without life, the planet's chemical state, energy balance and history cannot be understood. In *Revolutions That Made the Earth*, Earth-systems scientists Tim Lenton and Andrew Watson describe the shaping of our planet by life, combining evolutionary biology and geochemistry to do so.

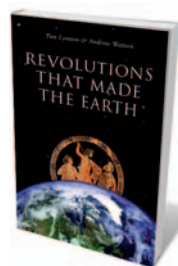
Lenton and Watson relate how life forms have co-evolved with the planet's environment over billions of years, and suggest explanations for some of the mysteries in the geological record. They focus on two major events that transformed Earth. Both were caused by innovations in the biological realm that were difficult to achieve in evolutionary terms, and both boosted the use of energy and recycling of materials crucial to life.

The first event is the oxygen revolution, triggered by the emergence of oxygenic photosynthesis more than 2.7 billion years ago, possibly in cyanobacteria. That led to the 'great oxidation' 300 million years later, when Earth's atmosphere retained oxygen

for the first time — the greatest chemical alteration in the planet's history.

The second is the complexity revolution that began around 1 billion years ago with the rapid rise of eukaryotic life — cells with nuclei that now make up all plants and animals. It allowed the conquest of land and the advance of complex multicellular life forms in the Cambrian explosion 542 million years ago, with profound geochemical consequences for the flows of carbon dioxide, oxygen and phosphorus that led to the planet we inhabit today.

Lenton and Watson are propelled by more than intellectual curiosity. By understanding previous planetary revolutions, they propose that we might be able to draw cautious conclusions about the upheavals that humans are



Revolutions That Made the Earth
TIM LENTON AND ANDREW WATSON
Oxford University Press: 2011. 448 pp.
£29.95, \$52.95

currently causing on Earth by melting its ice caps and deforesting its surface. They suggest that our ability to communicate through language, and to create extensive cultures through the networking of our messages, can be seen as the unfolding of another rare and hard-won revolution in the history of life — a shift in global consciousness.

Earlier transitions, the authors show, were followed by periods of turmoil in the Earth system, caused by instabilities inherent to its complex behaviour. There was no guarantee that life would survive these fluctuations: runaway effects twice locked Earth into global glaciations when greenhouse-gas concentrations plummeted. They could have turned Earth into a heated desert when they ended. But over tens of millions of years, feedbacks established chemical cycles of carbon and oxygen that once more kept Earth within a corridor of habitability. Crises thus arose separately from meteorite impacts and volcanic activity.

The message for our time is that the current planetary state could equally be destabilized if human activity causes critical Earth-system thresholds to be passed. With enough warning and insight, we might avoid or limit the damaging changes, or adapt to them rather than simply suffering the consequences. But the current lack of progress in mitigating climate change and preserving complex ecosystems is not promising.

Lenton and Watson highlight two concerns.

► NATURE.COM
Andrew Watson reviews Lovelock's *Revenge of Gaia*:
go.nature.com/gsw5qz

First, when Earth is operating near a tipping point, small changes in the driving factors can tip it rapidly from one state to another. We do not know how much the radiation balance of the atmosphere, or the ecological richness of ecosystems, or the chemical state of the oceans can be altered before cascades of effects are set off. Human activity could trigger a switch in important Earth-system processes even though the underlying change is gradual.

Their second observation is that former revolutions left life using more energy, with increased flows of materials from the environment, recycling of essential elements through life and increased networking of biological information. The authors extrapolate that for human societies to remain viable while using large amounts of energy to process large volumes of matter and maintain densely populated landscapes — as we do today using fossil fuels — we will have to boost the rate of recycling. The consciousness revolution must be followed by a recycling revolution.

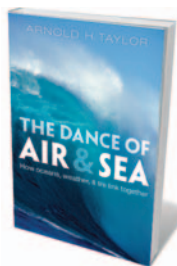
Lenton and Watson's thought-provoking book is the latest in a distinguished line of works that have altered our perception of the planet. Russian-Ukrainian geochemist Vladimir Vernadsky first discussed the deep involvement of life in planetary chemistry in his 1926 book *Biosfera* (Biosphere). In *Gaia* (Oxford University Press, 1979), James Lovelock brought the self-stabilizing mechanisms of life into view by seeing the planet as a partially self-regulating, living whole. And Hans Joachim Schellnhuber's book-length chapter in *Earth System Analysis* (Springer, 1998) laid out a blueprint for a scientific discipline concerned with the interplay of social and environmental dynamics.

The inventor R. Buckminster Fuller proposed in the 1950s to construct a giant globe of light bulbs to display data sets of global phenomena. His idea was that an external, physical realization of world mechanisms would engage our minds and change our ecological perspective. He was convinced that if we did not become an active, consciously reflective part of planetary evolution, we would become an unwitting, passive part of it.

Lenton and Watson's *Earth* — shaped by life, an escapee of close calls and again on a perilous course in the wake of biological innovation — is a similarly novel depiction. The Copernican revolution had no direct effect on people's everyday lives, but it ushered in a new era of self-consciousness. Perceiving Earth as a complex system of co-evolution similarly alters our outlook. ■

Wolfgang Lucht is the Alexander von Humboldt Chair in Sustainability Science at the Potsdam Institute for Climate Impact Research and at Humboldt University Berlin, D-14412 Potsdam, Germany. e-mail: wolfgang.lucht@pik-potsdam.de

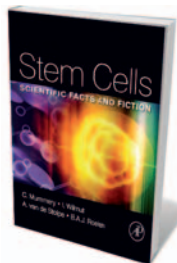
Books in brief



The Dance of Air and Sea: How Oceans, Weather, and Life Link Together

Arnold H. Taylor OXFORD UNIVERSITY PRESS 288 pp. £16.99 (2011)

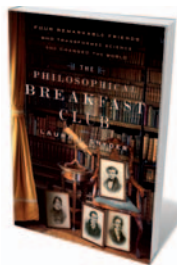
Recurring weather patterns such as El Niño and the North Atlantic Oscillation underlie many changes in biological systems. Marine scientist Arnold Taylor explains how great global circulations come about through ties between air and sea, and how they impinge on life. He describes how they were discovered, through tracking drifting wrecks, plastic toys and other flotsam. And he cautions that, despite the cycles' regularity, their complex and nonlinear origins make predictions of their future behaviour difficult.



Stem Cells: Scientific Facts and Fiction

Christine Mummery, Ian Wilmut, Anja Van De Stolpe & Bernard Roelen ACADEMIC PRESS 324 pp. \$79.95 (2011)

Stem-cell research holds great promise but is mired in controversy. Although our understanding of regenerative tissues has increased in recent decades, clinical applications of stem cells for treating chronic diseases such as diabetes and Parkinson's have proved hard to develop. Ethical issues also surround the use of embryonic stem cells. Translated into English from Dutch, this balanced book aims to set out all sides of the debates for both scientist and non-scientist readers, from researchers to patient advocates.



The Philosophical Breakfast Club: Four Remarkable Friends Who Transformed Science and Changed the World

Laura Snyder BROADWAY 448 pp. \$27 (2011)

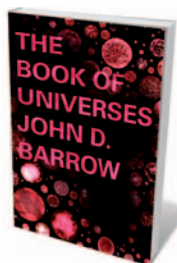
Cambridge University has schooled many great scientists. Historian Laura Snyder explores the friendships between four men who met there in the 1810s: Charles Babbage, inventor of the computer, astronomer John Herschel, crystallographer William Whewell and economist Richard Jones. Inspired by their seventeenth-century forebears, including Francis Bacon, they founded a breakfast club, where they plotted to revolutionize science. Drawing on their correspondence, Snyder describes how they did just that.



An Optimist's Tour of the Future: One Curious Man Sets Out to Answer "What's Next?"

Mark Stevenson PENGUIN 384 pp. \$26 (2011)

Science communicator and stand-up comedian Mark Stevenson travels the globe to interview top thinkers about what the future holds. From Harvard University geneticist George Church to the President of the Maldives, Mohamed Nasheed, he examines trends in personal medicine, robotics, nanotechnology and climate-change mitigation. He concludes that the future will be humane despite the growth of technology. People will adapt to change, as they always have done, and networks will replace hierarchies.



The Book of Universes

John D. Barrow BODLEY HEAD 368 pp. £20 (2011)

Mathematician John Barrow ponders the myriad types of universe — expanding, spinning, with and without life — that fall out of Albert Einstein's equations of general relativity. From Aristotle's spherical cosmos to the latest ideas about parallel universes, Barrow eruditely describes the many possibilities that astronomers are rushing to test through observations. Yet he admits that ultimately, even though our knowledge has come a long way, the Universe is still baffling.



Novelist Jane Austen wrote about the moral philosophy of sensibility in 1811, at the peak of its popularity.

HISTORY

Science and sensibility

Mechanical world views were replaced by more sensory beliefs after the rise of the novel, finds **George Rousseau**.

Mechanical philosophies of nature flowered in the seventeenth and eighteenth centuries when Enlightenment beliefs, grounded in reason and empirical observation, came to the fore. The whole Universe, it was thought, could be explained in terms of matter undergoing motions that were governed by physical laws. Most scientists today are aware of the heritage of mechanism in Isaac Newton's theories about force and matter, motion and velocity, cause and effect. Less familiar is the philosophical phase that followed — sensibility,

the view of humans as organic creatures, incapable of reduction to the sum of their mechanical parts, especially in the affective, moral and political realms.

In this weighty book, part of his multi-volume history of philosophy, Stephen Gaukroger explains how the philosophies of mechanism collapsed over eight decades, to be replaced by a more sensory view of nature. Gaukroger thoroughly comprehends the deep layers of science on which he pronounces. But by devoting the vast majority of the book to mechanism, he renders his treatment of

sensibility too short to fully document the shift from one mind-set to another. Readers will have to wait for the next volume to make sense of this one.

Explaining this transition is a tall order because of its complexity. Mechanism was never a single set of principles about machine-like systems. It comprised an array of disparate beliefs, experiences and practices that were followed in far-flung places and presided over by its principal architects: René Descartes, Thomas Hobbes, John Locke and Newton.

Gaukroger, who has worked in this field since the late 1970s, deftly describes mechanism's variety. He accomplishes his goal brilliantly through a slow narrative pace that is used for all but the last part of his book. If only he had laboured as unhurriedly over sensibility. Gaukroger's riposte to this will be that sensibility, as a concept, did not infiltrate Western thought until the mid-eighteenth century; that is, at the end of his book's chronological span. He is right, but without the next (unpublished) volume in hand, few readers will understand the concept's hydra-headed complexities. Sensibility was as varied a set of beliefs and practices as mechanism.

Sensibility is a vague term. Is it a scientific concept, a state of mind or a description of the senses of particular individuals? Writing at the peak of such philosophy, Jane Austen teased out the strengths and weaknesses of the word in her novel *Sense and Sensibility* (1811). Two generations earlier, Denis Diderot gave it a bipartite definition in his *Encyclopédie*, assembled in the 1750s and 1760s, as "the soul's disposition" and "the sensitive principle of life". Gaukroger's description is far stronger: "Sensibility", he writes, "allows connections to be made between natural-philosophical and moral, political, and psychological theories in a new way, shaping a new field of the moral sciences".

Gaukroger argues rightly that neither science nor modernity can be understood without comprehending why mechanism collapsed and sensibility replaced it. But with his scant treatment, like others who have tried, he does not have the final say. His explanation centres around fracture and fragmentation:

mechanism was not depleted but splintered, he argues; it unravelled and succumbed to new practices that placed greater emphasis on



The Collapse of Mechanism and the Rise of Sensibility: Science and the Shaping of Modernity 1680-1760

STEPHEN GAUKROGER
Oxford University Press:
2010. 505 pp. £35

➔ **NATURE.COM**
Gaukroger's *The Emergence of a Scientific Culture*:
go.nature.com/vab4zi

humans than on physical matter. In an age of increasing professionalization, especially among botanists, chemists, geologists and physicians, philosophical approaches diverged.

As mechanism was fracturing, Gaukroger notes, a new literary form was emerging that would develop into the novel. He writes, "It is internal psychological exploration, in the form of refined sensibility, rather than external events that have become the focus of the novel". It is true that sensibility needed these new forms of prose, and that it underlies our cognitive life far more than mechanism does. But such pronouncements are too few to explain why sensibility toppled centuries of mechanical thought.

Imaginative literature, later codified as 'Romantic', also drove nails into mechanism's coffin by postulating that matter was more complex than the mechanical natural philosophers thought. A human is not a mere machine; a fly is much harder to study than a pebble. By focusing on human nature rather than physical matter, the language of the new literature helped to alter the way scientists conceived their models, and enabled modernity to commence its work.

In this sense, the 1760s was a watershed decade — except that sensibility had already reared its head in French imaginative literature a century earlier. Just when you think Gaukroger has sent mechanism to its grave, he weaves in broader elements of history that further complicate the evolving drama of mechanism's failures. This is to be expected — intellectual history rarely evolves in orderly ways. Its fault lines are unkempt and its major paradigms overlap. It is to Gaukroger's credit that he avoids facile explanation.

Massive shifts in the history of knowledge are difficult to chart. The late historian of science Thomas Kuhn resorted to the intricate concepts of 'paradigm shifts', whereas philosopher Michel Foucault elaborated the acceptance of certain ideas at certain times in his 'epistemes'. But historians and scientists have different expectations of what constitutes adequate context. In my view, Gaukroger has soft-pedalled the socio-cultural aspects of sensibility. His approach sheds new light on mechanism's decline, but less on its organic replacement. I hope that a subsequent volume will rectify the imbalance. ■

George Rousseau is professor of history and co-director of the Centre for the History of Childhood, University of Oxford, Oxford, OX1 4AU, UK. He is author of *Nervous Acts*.
e-mail: george.rousseau@magd.ox.ac.uk

CHEMISTRY

Enigmatic elements

A Philadelphia exhibition is a playful celebration of the periodic table, reports **Katharine Sanderson**.

Elements, the fundamentals of chemistry, have an almost mystical symbolism. The astounding concept that all matter and all life comes from the coupling of energy with this ordered list of atoms unifies an exhibition to celebrate the International Year of Chemistry 2011. *Elemental Matters*, which opened this month at the Chemical Heritage Foundation in Philadelphia, Pennsylvania, and runs throughout 2011, features the work of seven artists known for tackling chemistry themes.

Canadian artist David Clark jokes that he was inspired to create pieces based on the periodic table because the symbols for the consecutive elements chlorine, argon and potassium spell out his surname. His work focuses on the structure of the table rather than its chemical contents. In *I Don't Think You Understand the Way I Feel About the Stove* (2000; pictured) — borrowing from the words of *Stove*, a song by Canadian indie rock band Eric's Trip — he replaces the chemical symbols with 118 identical rusty electric-stove heating elements. "By collecting objects that are all the same, it emphasizes the table's meaning as a sign," he says.

In *Braille* (2000), Clark reorganizes the elemental symbols into another familiar

Elemental Matters: Artists Imagine Chemistry
Chemical Heritage Foundation, Philadelphia, Pennsylvania.
Until 16 December 2011.

chart, the eye test. The letters, also translated into Braille, shrink in size with each descending line. "The Braille fades into nothingness," says Clark. "They are like atoms, moving

beyond our touch." His intent is to remind us that the shape of the periodic table, even without the information it usually holds, has become iconic.

The chemical ingredients of the human body are decoupled in New York artist Dove Bradshaw's *Self Interest* (1999), a series of 58 flasks mounted in a glass case. Each container holds a sample of an element scaled to the amount found in a person weighing 45 kilograms. The flasks containing trace elements such as yttrium, thorium or beryllium are only as big as three pinheads, notes Bradshaw. The piece explores whether our existence can be stripped down to material constituents.

The elemental forces of nature, and the chemical changes tied up with those forces, also find a place in this exhibition. In Bradshaw's *Waterstone* (1996), for example, a funnel drips water at a steady rate onto a limestone block. Running since the mid-1990s, the trickle of seven drops every minute has eroded a small dent in the carbonate.

The displayed works include more literal takes on the elements — in sound as well as vision. *Elements in Descending Order of Creation from Collapsing Stars*, by Oregon-based composer Susan Alexander, transposes the vibrational frequencies of atoms down by 14 octaves into the audible range. Her musical 'scale' runs from hydrogen, helium, carbon, nitrogen and oxygen to silicon, phosphorus and sulphur.

The show also includes collaborative works, such as a giant printed collage of the 118 elements of the periodic table interpreted by 97 artists, with each print in its conventional place.

Chemists usually think of the periodic table in abstract terms, as a reference book or an aid to their research. These artworks remind us of the mystery that the elements can also evoke. ■

Katharine Sanderson is a writer based in France.



Stove elements replace chemical ones in David Clark's artwork inspired by the periodic table.



Q&A Roman Kaiser

Perfumes preserved

Roman Kaiser, a chemist at the Givaudan Research Centre in Dübendorf, Switzerland, has recreated the scents of hundreds of rare and endangered plants in his laboratory, some of which he describes in his latest book. Here he explains how he preserves the smell of disappearing flora.

Why study the scents of rare and endangered species?

A third of all flowering plants are expected to become extinct by the middle of this century. Plants are essential for life, so preserving their biodiversity is as important as preventing climate change. My book *Scent of the Vanishing Flora* collects olfactory descriptions for 267 rare and endangered plants (selected from 500 I've studied), a few of which no longer exist in nature. These descriptions will allow experts to recreate them. I've also prepared a small number of boxes containing 60 scent reconstructions, as exemplars. When these plants no longer exist, we will still be able to smell them.

Where have you travelled to collect scents?

As a child I wandered in the forests around our village in Switzerland, so I started with the black vanilla orchid (*Nigritella nigra*) that grows in those mountains. In the 1980s, I visited the Mediterranean countries and the rainforests of Indonesia in Java and Bali. Then the expeditions became bigger. In 1996, I traversed French Guiana with a zeppelin-like dirigible to sample rare and endangered species from the canopy of the rainforest, and I later undertook similar expeditions in Gabon and Madagascar. In 2001, I explored the mountain mist forests of Papua New Guinea, which are rich in species not yet described. I have also visited a

plateau close to the western coast of South Africa where you can find a quarter of all bulbous plants that exist.

Why do flowers have a smell?

To attract pollinators. Some flowering plants release their scents at night to attract moths — their fragrances are often similar to jasmine, orange flower or gardenia. Others use mimicry and deception. In my lab I have a species of *Stapelia* from South Africa that, when in bloom, uses a smell like decaying meat to attract carrion flies. In the Andes there are orchids that emit a mushroom-like scent to attract female fungus flies. There is an orchid named after me that attracts male euglossine bees, who visit not for nectar but to use the fragrance to attract female bees. There are many more strategies that flowers apply to get pollinated.

How has the study of natural scents developed?

Until the mid-nineteenth century, we had only natural extracts from plants and animal secretions. With the dawn of modern organic chemistry we could investigate the olfactory compounds within these extracts. By 1970, we knew of hundreds of extracts and more than 1,000 synthetic fragrance and flavour compounds. But this didn't even begin to cover the huge range of attractive natural scents. Fortunately, around this time, modern electronics

Scent of the Vanishing Flora
ROMAN KAISER
Wiley: 2010/2011.
400 pp. £105/\$155

gave us sensitive devices with which to non-destructively collect micro-samples of scents directly from living plants, preserving their quality as perceived by the human nose. This avoids influencing the original scent, which was a problem with earlier isolation methods such as distillation.

How do you capture a fragrance in the wild?

We place a small absorption trap close to the scent source (usually a flower) and pump the fragrant air through it. The scent molecules are absorbed over minutes or hours. You need no more than 20–200 micrograms for a laboratory investigation using gas chromatography and mass spectrometry. Then you can reconstitute the natural scents from synthetic scent compounds.

How does smell compare to the other senses?

Visual and acoustic perception can be described mathematically, but you can only describe a new smell by comparing it to your existing scent experiences. Novelty in a scent does not manifest itself in a direct way: you will feel uncertain and have trouble describing it. And then you have to study it.

What makes a smell good or bad?

In certain cases, concentration has an enormous role. One powerful olfactory chemical is 4-methyl-4-sulphanylpentan-2-one, probably the most potent odorant in the world. Less than 0.0001% of it in a solution gives the attractive cassis notes in sauvignon blanc wine. But in 1993 we found a Swiss orchid in which the same chemical is present at 0.01%. It has the scent of tomcat urine.

What are the least pleasant odours you've worked with?

I'm sometimes asked by museums to reproduce animal scents, such as the smell of cattle in a barn. I was once asked to make the odour of a skunk. I made a diluted version of it, but the people at the museum were not careful. A dispenser containing just 10 millilitres of a 0.001% solution of the reconstituted skunk scent was spilt, and it contaminated a large part of the building.

What does the future hold?

As analytical systems become more sensitive, we will know more about the composition of natural scents. In the future it will be easier to synthesize organic chemicals with milder methods and less waste. In theory, the process of analysing and recreating a natural odour could be almost automated. But the human nose will still be the most sensitive detector. ■

INTERVIEW BY JASCHA HOFFMAN

CORRESPONDENCE

Differing opinion, not misconduct

It greatly concerns us that the reputation of one of our faculty, cancer-vaccine researcher Nina Bhardwaj, was called publicly into question through your premature publication of fired researcher David O'Neill's allegations of research misconduct (*Nature* 467, 260; 2010), particularly as the New York University School of Medicine (NYU) was not then in a position to comment. The investigation by the NYU inquiry committee is now complete and Bhardwaj's reputation is restored.

The committee investigated each of O'Neill's allegations, reviewing relevant documents and drafts of manuscripts, interviewing the individuals involved, and engaging an outside expert statistician to review the allegations and documents. Neither the committee nor the independent expert found any credible evidence to support the allegations (for details, see go.nature.com/gxhfm1).

Before O'Neill's employment was terminated (which was not in retaliation for reporting alleged misconduct as he said), the other co-authors of the manuscript about which he had complained convened to address his disagreement with the method of statistical analysis used. The dispute was resolved by including three different statistical methodologies in the manuscript, which reported on the outcome of an exploratory clinical trial. Each of these came to the same conclusion — namely, that the dendritic-cell vaccine developed by Bhardwaj is inferior to vaccine delivered by mineral-oil adjuvant.

The committee found no credible evidence that Bhardwaj was motivated, financially or otherwise, to “spin” the results of the trial to promote the dendritic-cell vaccine. Along

with the independent expert, they concluded that the statistical tests were selected by the director of NYU's biostatistics division, not by Bhardwaj, and that the disagreement constituted a difference of opinion, not research misconduct. This difference arose because Bhardwaj and the other co-authors gave greater weight to the opinion of the statistical experts than to O'Neill on the matter of statistical method.

Bhardwaj's standing has therefore been upheld as a valued member of our faculty, as a committed mentor to junior faculty members and as a highly esteemed researcher in the cancer-vaccine field.

Steven B. Abramson *New York University School of Medicine, USA.*
stevenb.abramson@nyumc.org

Stuttering studies support treatment

We take issue with some of Peter Howell's statements about stuttering (*Nature* 470, 7; 2011).

Most of these relate to the Lidcombe programme for reducing stuttering, which we developed with colleagues at the Australian Stuttering Research Centre at the University of Sydney. Statements about the efficacy of the treatment are based on research data, not belief — as Howell implies.

We dispute his claim that there are insufficient data to support firm conclusions about the treatment. There are more than ten peer-reviewed journal reports evaluating the treatment positively, including a randomized controlled trial of 54 children published in 2005 in the *British Medical Journal*, and an independent, randomized study of 46 German children. A meta-analysis of randomized, controlled evidence for the treatment of 145 children, to be published in the *Handbook of Evidence-Based Practice in Clinical Psychology*

(Wiley, in the press), shows that those who receive it have 7–8 times greater odds of attaining satisfactory clinical outcomes than those who do not.

We are insulted that Howell likens the procedures used in the treatment to training a dog, given the extensive literature about the use of verbal operant procedures (such as reward and punishment) in many aspects of clinical psychology and education. His view is that the results of our clinical trial could be skewed because we consider that repeated words (not just syllables) can constitute stuttering. The World Health Organization makes no such exclusion.

With regard to auditory feedback devices, we would like to point out that a clinical trial has shown no treatment effect in real-life situations (R. Pollard *et al. J. Speech Lang. Hear. Res.* 52, 516–533; 2009). It seems to us that Howell also underestimates the incidence of stuttering. A prospective study published in 2009 in *Pediatrics* of a cohort of 1,619 children (ascertained before the onset of stuttering) reported a cumulative incidence of stuttering at age 3 years of 8.5%.

A prospective longitudinal study of 147 stuttering children in 1999 also showed that the severity of stuttering in children younger than 6 years did not predict lifelong stuttering. We therefore disagree with Howell's view that parents should be told whether a child's recovery is predictable.

Many reports show that stuttering typically starts at 2–4 years of age, and most children who are destined to recover naturally will have done so by age 8. Therefore, we also question Howell's assertion that he can predict stuttering recovery when children are 8 years old.

Mark Onslow, Ann Packman *Australian Stuttering Research Centre, The University of Sydney, Australia.*
mark.onslow@sydney.edu.au

Peter Howell replies:

A comprehensive list of technical references to support my arguments (*Nature* 470, 7; 2011) is available at go.nature.com/pzlmgw. These relate to statistical issues and the Lidcombe programme; to why I hold that single-syllable whole-word repetitions should not be considered symptoms of stuttering; to my position that use of whole-word repetitions overestimates incidence rate; to why I contend that severity predicts risk of stuttering persistence; and to my view that the severity model can be used to predict persistence of stuttering beyond the age of eight (P. Howell and S. Davis *J. Dev. Behav. Pediatr.*, in the press).

The reason why I pointed out that operant work originated in animal studies was to indicate that clinical studies could also benefit from examination of recent advances in animal operant work.

Peter Howell *University College London, UK.*
p.howell@ucl.ac.uk

Clarifying relatives and ancestors

You twice erroneously describe the extinct woolly mammoth and American mastodon as ancient “ancestors” of today's elephants (see go.nature.com/q7fvk6 and go.nature.com/91a4zy).

They are actually fossil relatives of today's elephant species, as is made clear in the original research paper (N. Rohland *et al. PLoS Biol.* 8, e1000564; 2010) on which the two items report. With the extensive range of expertise at the disposal of your journal, I am surprised at these slips.

Walter Shearer *Hartsdale, New York, USA.*
wshearer@optonline.net

REUTERS/P. ANDREWS



CENTRE: D. PRICE/ALAMY; RIGHT: IMAGE SOURCE/GETTY

Figure 1 | Post-traumatic stress disorder (PTSD). Going to war, and experiencing severe physical assault and injury are all situations that can lead to PTSD.

PSYCHIATRY

A molecular shield from trauma

Why do some people develop post-traumatic stress disorder, but others emerge from a horrific event relatively unscathed? A molecule involved in orchestrating the brain's response to stress may hold the key to this difference. [SEE ARTICLE P.492](#)

MURRAY B. STEIN

Exposure to traumatic, potentially life-threatening events can change forever the way an individual views themselves and the world around them. In mental-health parlance, when one can't stop thinking (or dreaming) about the event, and can't restrain the fear response and resume everyday functioning, post-traumatic stress disorder (PTSD) has taken hold. On page 492 of this issue, Ressler *et al.*¹ report that PACAP — a neuropeptide involved in coordinating the brain's response to stress — and *ADCYAP1R1*, the gene that encodes PACAP's receptor PAC1, are pivotal in determining whether an individual develops this disorder.

A range of experiences can cause PTSD. Examples include being the victim of physical or sexual assault (often on a chronic basis, as in the case of domestic violence); experiencing severe physical injury, perhaps as a result of a car crash; and being a combatant in armed conflicts (Fig. 1). Indeed, more than 10% of US Army soldiers return from deployment suffering from PTSD or major depression², both of which increase the risk of suicide.

As testimony to the strength and resilience of the human spirit — and bodily stress-response systems — most people exposed to horrific events rebound emotionally³. Others do not recover (probably 10–20%, depending on the severity of, and the length of exposure to, the trauma), and develop chronic PTSD. Ressler *et al.* set out to investigate what, at the molecular level, determines

this difference between individuals.

The authors¹ studied a sample of urban primary-care patients whom they had selected for being at high risk of PTSD: the patients had histories of extensive exposure to traumatic situations, such as a serious accident, being attacked with a weapon, or having been maltreated as a child. Among the subjects, peripheral blood levels of PACAP correlated positively with the extent of post-traumatic stress symptoms, but only in women. Moreover, a common genetic variant in *ADCYAP1R1* was associated with PTSD. This finding, which Ressler and colleagues¹ replicated in a second sample of patients exposed to trauma, was once again limited to women. Intriguingly, the variant within *ADCYAP1R1* resides in a putative oestrogen response element. This observation provocatively offers a possible explanation for why women might be more likely than men to develop PTSD⁴, even when exposed to similar types and 'doses' of trauma.

Maltreatment during childhood is a strong risk factor for subsequent psychological disorders, including PTSD⁵. Such environmental experiences can result in epigenetic changes — specifically, DNA modification through methylation — which alter gene transcription⁶. In mice, for instance, variation in the extent of methylation of *Crf*, an essential gene for the stress response, helps to explain individual differences in the animals' behavioural response to chronic social stress⁷. Also, Ressler *et al.* show that the extent of methylation of a site in *ADCYAP1R1* is associated with PTSD in their

patients — a finding that was not restricted to women. Together, these observations strongly suggest that one way in which traumatic stress leaves a cumulative and enduring stamp on the human genome and psyche is methylation of genes that are essential for the stress response.

The neurobiology of PTSD is poorly understood. Brain imaging studies, however, indicate that a reduced hippocampal volume is a risk factor⁸. Also, abnormal functioning of the fear-related circuitry, which includes the brain's amygdala and insula, is a possible biomarker for PTSD⁹.

Such limited advances in understanding the pathophysiology of PTSD have failed to translate into the development of new treatments. Current treatments for PTSD include specialized psychological therapies¹⁰ and drugs, such as antidepressant medications that block reuptake of the neurotransmitter serotonin¹¹. But these treatments are effective for only some individuals. Moreover, there is at present no easy way of telling who will benefit from a certain treatment, thereby necessitating individual therapeutic trials for drugs — a practice that can be frustratingly slow and costly.

An earlier study¹² suggested that a particular 'signature' of individuals exposed to trauma (obtained through gene-expression profiling of their peripheral white blood cells) could predict who would develop PTSD four months later. Do Ressler and co-workers' findings, which shed light specifically on the PACAP–PAC1 pathway, bring us closer to obtaining a biomarker for PTSD? Will it be possible, for example, to measure PACAP levels in a

person's blood after they have been exposed to severe traumatic stress and determine whether they will develop PTSD? Or could sequencing influential variants within *ADCYAP1R1* determine who is at risk of developing this disorder?

The answer to all of these questions is a qualified maybe. First, however, it is essential that these results are replicated by other research groups, and in larger samples. Also, more must be learned about the relationship between PACAP levels in peripheral blood and cerebrospinal fluid, and, particularly, in the regions of the brain that modulate the fear response. Finally, greater molecular insight is needed into how the observed variant in *ADCYAP1R1* affects the risk of PTSD, and the extent to which its effects relate to those of other genetic loci — such as *FKBP5* and *5HTTLPR*, which have also been implicated^{5,13,14} in PTSD and in risk-resilience to extreme stress.

Even if PACAP is not it, this proof of principle study¹ shows that developing a reliable biomarker for PTSD is not a pipe dream. For example, it might be possible to develop a PTSD risk profile that takes into account psychosocial and biological factors, much like those available for predicting the risk of coronary heart disease. Armed with that information, investigators might be able to conduct clinical trials of new drugs — perhaps involving small molecules that modify the activity of the PACAP–PAC1 pathway — intended to enhance resilience to (and so prevent) the serious consequences of traumatic stress. ■

Murray B. Stein is in the Departments of Psychiatry and of Family & Preventive Medicine, University of California, San Diego, La Jolla, California 92093-0855, USA. He is also in the Psychiatry Service at the Veterans Affairs San Diego Healthcare System, San Diego.
e-mail: mstein@ucsd.edu

1. Ressler, K. J. *et al.* *Nature* **470**, 492–497 (2011).
2. Thomas, J. L. *et al.* *Arch. Gen. Psychiatry* **67**, 614–623 (2010).
3. Galea, S. *et al.* *N. Engl. J. Med.* **346**, 982–987 (2002).
4. Roberts, A. L., Gilman, S. E., Breslau, J., Breslau, N. & Koenen, K. C. *Psychol. Med.* **41**, 71–83 (2011).
5. Binder, E. B. *et al.* *J. Am. Med. Assoc.* **299**, 1291–1305 (2008).
6. Murgatroyd, C. *et al.* *Nature Neurosci.* **12**, 1559–1566 (2009).
7. Elliott, E., Ezra-Nevo, G., Regev, L., Neufeld-Cohen, A. & Chen, A. *Nature Neurosci.* **13**, 1351–1353 (2010).
8. Gilbertson, M. W. *et al.* *Nature Neurosci.* **5**, 1242–1247 (2002).
9. Etkin, A. & Wager, T. D. *Am. J. Psychiatry* **164**, 1476–1488 (2007).
10. Bisson, J. & Andrew, M. *Cochrane Database Syst. Rev.* CD003388 (2007).
11. Ravindran, L. N. & Stein, M. B. *Brain Res.* **1293**, 24–39 (2009).
12. Segman, R. H. *et al.* *Mol. Psychiatry* **10**, 500–513 (2005).
13. Xie, P. *et al.* *Neuropsychopharmacology* **35**, 1684–1692 (2010).
14. Stein, M. B., Campbell-Sills, L. & Gelernter, J. *Am. J. Med. Genet. B* **150**, 900–906 (2009).

SPECTROSCOPY

Unexpected interactions

Unpaired electrons can exert effects that allow interatomic contacts in molecules to be detected more easily using nuclear magnetic resonance. One such effect reveals unusual interactions between certain atoms in a protein.

IVANO BERTINI & CLAUDIO LUCHINAT

In folded proteins, it is quite easy to find places where the distance between the hydrogen atom in a carbon–hydrogen (C–H) bond and a sulphur atom in a nearby cysteine amino-acid residue is less than 4 ångströms. But no one had suggested that this might be evidence of a weak hydrogen-bond-like interaction — until now. Reporting in the *Journal of the American Chemical Society*, Westler *et al.*¹ describe unexpectedly large shifts of peaks associated with certain carbon nuclei in the nuclear magnetic resonance (NMR) spectrum of the protein rubredoxin. The shifts correspond to ¹³C nuclei that are close to sulphur atoms covalently bound to the iron atom in the protein's active site. The authors attribute the shifts to C–H...S–Fe interactions (where the dotted line indicates a non-covalent interaction and S–Fe indicates a covalent sulphur–iron bond), and conclude that these interactions may make a small but non-negligible contribution to the stability of the protein's three-dimensional structure.

Rubredoxin is a small protein that takes part in electron-transfer processes in certain microorganisms (bacteria and archaea). The active site of the protein contains an iron ion bound by four negatively charged sulphur atoms, in which the iron cycles between the Fe²⁺ and Fe³⁺ oxidation states as the protein performs its functions². Westler *et al.*¹ compared the NMR signals of ¹³C nuclei lying close to sulphur atoms in the chemically reduced form of rubredoxin with the signals of the same nuclei in an artificial version of the protein in which the iron ion had been replaced with a zinc ion, Zn²⁺. They observed sizeable shifts in the positions of these signals, which they attributed to the presence of four unpaired (paramagnetic) electrons in Fe²⁺; unpaired electrons are not present in Zn²⁺.

Shifts in NMR signals associated with unpaired electrons are called hyperfine shifts, and can be caused by effects transmitted through bonds or through space³. The carbon nuclei of interest in rubredoxin are separated from the nearby iron atoms by many covalent bonds, suggesting that the hyperfine shifts observed by Westler *et al.*¹ are not caused by effects transmitted through those bonds. But the authors convincingly argue that the shifts cannot be due to through-space effects, because the iron is too far away from the

carbon atoms for this to be possible. Instead, they suggest that the shifts are caused by the delocalization of a tiny fraction of the unpaired Fe²⁺ electrons onto the carbon nuclei. But how could this happen, if classical through-bond and through-space effects have been ruled out?

The authors propose that the shifts are caused by C–H...S–Fe interactions, in which a hydrogen atom acts as a bridge between an iron-bound sulphur atom and a carbon atom (Fig. 1). In other words, the hyperfine shifts are through-bond effects that act through hydrogen-bond-like C–H...S–Fe interactions. If so, the size of each shift — and therefore the

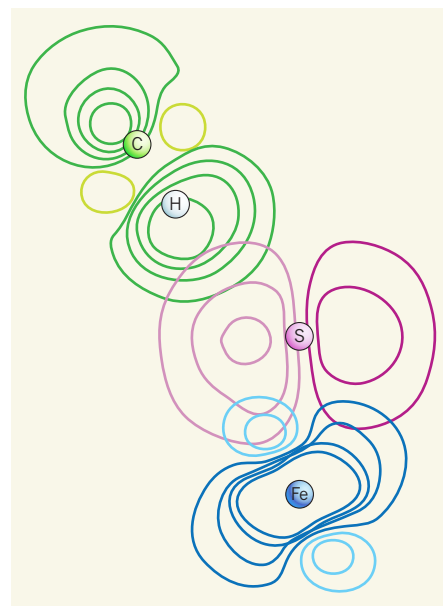


Figure 1 | Overlapping orbitals in a through-bond interaction. On the basis of evidence obtained from NMR spectroscopy, Westler *et al.*¹ propose that hydrogen-bond-like interactions exist between certain carbon–hydrogen bonds and iron-bound sulphur atoms in the protein rubredoxin. The molecular and atomic orbitals associated with one of these interactions are depicted here. The overlap of a molecular orbital (green) on the C–H bond with orbitals on the sulphur and iron atoms (sulphur orbital, purple; iron orbital, blue) allows a tiny fraction of the unpaired electrons on the iron to be delocalized onto the carbon atom through the C–H bond. Spheres indicate the centres of the atoms; contour lines indicate different amounts of electron-spin density. For each orbital, the phase of the wavefunction of different lobes is indicated by different shades of the same colour. (Graphic based on Fig. 6 of ref. 1.)

amount of delocalization of unpaired electrons onto each carbon nucleus — would depend on the angle made by the interacting carbon, hydrogen and sulphur atoms. Westler *et al.* find that this is indeed the case.

To further bolster their argument, the authors modelled the rubredoxin system using density functional theory, a quantum-mechanical approach that is particularly suitable for revealing tiny electron-delocalization effects⁴. Their calculations confirmed that substantial hyperfine shifts (of more than 100 parts per million, p.p.m.) for ¹³C nuclei in C–H bonds occur when the H...S distance is 3 Å, provided that the C–H...S angle is at least 130°. The authors also calculated that the energy of the interaction between a hydrogen and a sulphur atom is less than 1,000 joules per mole — about one-tenth of that of a normal hydrogen bond (for example, for the N–H...S hydrogen bonds that are also observed in this system⁵), and consistent with what would be expected for this type of interaction.

A noteworthy aspect of Westler and colleagues' study is that dramatic effects in the NMR spectrum of rubredoxin were caused by a tiny amount of delocalization of paramagnetic electrons. We estimate that only about 0.01% of the iron's unpaired-electron density is delocalized onto the ¹³C nucleus in each C–H...S–Fe interaction, yet this provides an NMR shift of tens of p.p.m., a huge amount. The authors' study also backs up earlier NMR observations^{6,7} of through-bond effects between the cations and anions in certain salt solutions.

Somewhat paradoxically, hyperfine shifts transmitted through strong covalent bonds can also reveal tiny effects, but of a different nature. This happens in a class of proteins related to the rubredoxins — the iron–sulphur proteins, which contain four iron and four sulphide (S²⁻) ions arranged in a cluster. Here, hyperfine shifts of the hydrogens of iron-bound cysteines are transmitted from iron through the covalent bonds that connect the intervening sulphur and carbon atoms (from Fe to H in the Fe–S–C–H covalent-bond system). These shifts are associated with tiny variations in the magnetic moment of electrons in the iron ions that occur when the cluster releases or accepts an electron as the protein performs its biological function. From these shifts, one can identify which iron ions in the cluster are more prone to be reduced or oxidized⁸.

Westler and colleagues' discovery¹ adds to the catalogue of paramagnetic phenomena in NMR that are due to tiny effects, some of which are not yet theoretically understood. For example, in proteins that contain paramagnetic metals, shifts arising from through-space interactions should become attenuated with the cube of the distance from the metal³. Instead, especially for iron-containing haem proteins, nuclei in protein 'backbones' experience significant paramagnetic shifts even at large distances

from the metal^{9–11}, where the cube rule would predict negligible effects. In the light of Westler and colleagues' work, it is tempting to attribute these shifts to through-bond effects, caused by unpaired-electron delocalization mechanisms through the whole protein (which in turn would have important implications for electron-transfer phenomena¹²).

Similarly, in proteins containing a paramagnetic metal ion, the enhancement of the relaxation rates of nuclei (the rate at which nuclear magnetization returns to equilibrium from the non-equilibrium state induced during NMR experiments), caused by through-space electron–nuclei interactions, is theoretically predicted to decay with the sixth power of the metal–nuclei distance³. In reality, this enhancement for ¹³C and ¹⁵N nuclei levels off beyond a certain distance, rather than following the power law, as if a tiny fraction of an unpaired electron (or electrons) were delocalized across the entire protein¹³. Westler and colleagues' work¹ therefore not only provides hints about the weak interatomic interactions that contribute to stabilizing protein folds, but also will stimulate many fresh reflections on the elusive structural features made detectable by NMR through amplification provided by paramagnetic effects. ■

PRION DISEASES

Infectivity versus toxicity

Prions are infectious proteins that can cause deadly diseases in mammals. Detailed measurements of infectivity suggest that there may be distinct infectious and toxic versions of this protein. [SEE LETTER P.540](#)

REED B. WICKNER

P rions are unique in being transmitted by an altered form of a cellular protein — PrP in mammals — without the need for RNA or DNA¹. The disease form converts the normal form into one of its own. Mammalian prion diseases are all fatal, and include Creutzfeldt–Jakob disease in humans, bovine spongiform encephalopathy in cows, scrapie in sheep and chronic wasting disease in elk and deer. On page 540 of this issue, Sandberg *et al.*² provide evidence for two distinct PrP populations: infectious and toxic.

In the central nervous system of infected individuals, PrP accumulates as filamentous polymers called amyloid. This material has been considered to be the infectious species, and its accumulation in crucial brain parts has been assumed to cause neuronal degeneration and death. Precise identification of

Ivano Bertini and Claudio Luchinat
are at the Magnetic Resonance Center and
Department of Chemistry, University of
Florence, 50019 Sesto Fiorentino, Italy.
e-mails: bertini@cerm.unifi.it;
luchinat@cerm.unifi.it

- Westler, W. M., Lin, I.-J., Perczel, A., Weinhold, F. & Markley, J. L. *J. Am. Chem. Soc.* **133**, 1310–1316 (2011).
- Min, T., Ergenekan, C. E., Eidsness, M. K., Ichiye, T. & Kang, C. *Protein Sci.* **10**, 613–621 (2001).
- Bertini, I., Luchinat, C. & Parigi, G. *Solution NMR of Paramagnetic Molecules* (Elsevier, 2001).
- Justino, L. L. G. *et al. Inorg. Chem.* **47**, 7317–7326 (2008).
- Lin, I.-J. *et al. J. Am. Chem. Soc.* **131**, 15555–15563 (2009).
- Brown, D. G. & Drago, R. S. *J. Am. Chem. Soc.* **92**, 1871–1875 (1970).
- Bertini, I., Luchinat, C. & Borghi, E. *Inorg. Chem.* **20**, 303–306 (1981).
- Bertini, I., Ciurli, S. & Luchinat, C. *Struct. Bonding* **83**, 1–53 (1995).
- Guiles, R. D. *et al. Biochemistry* **32**, 8329–8340 (1993).
- Boyd, J., Dobson, C. M., Morar, A. S., Williams, R. J. P. & Pielak, G. J. *J. Am. Chem. Soc.* **121**, 9247–9248 (1999).
- Bertini, I., Luchinat, C. & Turano, P. *J. Biol. Inorg. Chem.* **5**, 761–764 (2000).
- Dempsey, J. L., Winkler, J. R. & Gray, H. B. *Chem. Rev.* **110**, 7024–7039 (2010).
- Ma, L., Jørgensen, A.-M. M., Sørensen, G. O., Ulstrup, J. & Led, J. J. *J. Am. Chem. Soc.* **122**, 9473–9485 (2000).

the infectious and toxic species has been difficult, however, owing to the heterogeneity of amyloids in general (and of PrP in particular), the cumbersome and expensive nature of the infectivity assay, and the absence of a simple toxicity assay.

Recently, a cell-based assay for prion infectivity has been developed³ that is both cost- and time-effective. Sandberg *et al.*² used this advance to precisely measure prion infectivity as a function of time in mice expressing different levels of PrP. They infected three mouse lines with the same dose of mouse-adapted scrapie prions: a transgenic line overproducing PrP roughly eightfold, a normal strain and a hemizygous line expressing only half the normal amount of this protein. The authors find that, irrespective of the amount of PrP each strain produces, all reach the same plateau of infectivity — a level beyond which prions do not increase in concentration

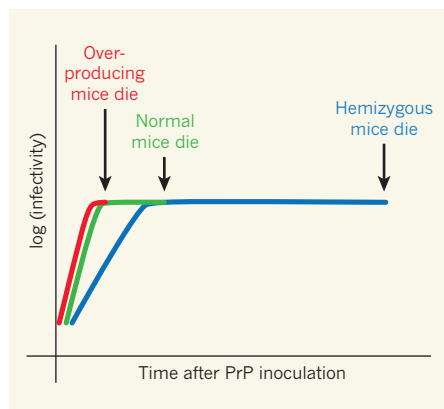


Figure 1 | Does PrP protein concentration matter? Sandberg *et al.*² show that, irrespective of how much PrP a mouse line produces normally, all lines reach the same plateau of infectivity after inoculation with infectious PrP. However, the amount of PrP an animal produces affects how soon it will die after the infectivity plateau is reached. (Adapted from ref. 1.)

(Fig. 1). This result indicates that there must be some limiting component, other than PrP monomers, that prevents the accumulation of infectious particles.

Although the PrP-overproducing mice died soon after the plateau was reached, the normal animals showed a long delay before death, and the hemizygous mice an even longer delay. Why does the same level of infectious material kill the overproducing mice much sooner than the other two strains? The authors infer that the toxic, killer PrP species must be distinct from the infectious species. This is unexpected and raises issues central to our understanding of prion diseases.

Although Sandberg and colleagues continuously measured the level of the infectious PrP, they measured only one value for the toxic species — that which kills the mice. They assumed that the same amount of the toxic species that kills the overproducing mice also kills the normal and the hemizygous mice. This might not be the case. The rapid accumulation of a smaller amount of the toxic species might kill the overproducing mice quickly, whereas mice slowly accumulating the toxic species might partially adapt, dying only when exposed to a greater amount of this species.

Toxicity is not well understood for any amyloid, and certainly not for the PrP amyloid in scrapie. However, only cells producing PrP are sensitive to the toxic species⁴: without the anchor (GPI) that secures PrP to the cell membrane and allows its internalization, prion propagation occurs, but there is little or no pathology⁵. In this sense, therefore, PrP is a 'receptor' for killing cells⁶. This would explain a direct relationship between toxicity and PrP expression level. Because PrP monomers assemble into amyloid on the cell surface or during internalization⁷, the amyloid might be gumming up some intracellular compartment,

such as the endosomes, thus causing cell death.

Because amyloid filaments grow from their ends, one might assume that each PrP amyloid filament is a potential infectious unit. However, filament toxicity might reflect adsorption or interaction of cellular components with the sides of filaments. Toxicity might, therefore, be proportional to the total length of filaments. Because chaperone proteins that cleave prion filaments are likely to be constantly active, the few filaments in hemizygous mice would be chopped into small pieces, the infectivity per mass of which would be high. By contrast, the PrP-overproducing mice would accumulate long filaments of great toxicity but lower infectivity per mass.

To identify the putative toxic form of PrP that Sandberg and colleagues² posit, it will be necessary to develop a rapid toxicity assay that shows the requirement for endogenous PrP expression for cells to be sensitive to the toxic effects, as is observed *in vivo*. Nonetheless, Sandberg *et al.* have shown, with a

simple experiment, that the infectious and toxic particles are not the same. Their report imposes several constraints on any explanation of this devastating group of diseases. ■

Reed B. Wickner is in the Laboratory of Biochemistry and Genetics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0830, USA. e-mail: wickner@helix.nih.gov

1. Caughey, B., Baron, G. S., Chesebro, B. & Jeffrey, M. *Annu. Rev. Biochem.* **78**, 177–204 (2009).
2. Sandberg, M. K., Al-Doujaily, H., Sharps, B., Clarke, A. R. & Collinge, J. *Nature* **470**, 540–542 (2011).
3. Klöhn, P.-C., Stoltze, L., Flechsig, E., Enari, M. & Weissmann, C. *Proc. Natl Acad. Sci. USA* **100**, 11666–11671 (2003).
4. Brandner, S. *et al. Nature* **379**, 339–343 (1996).
5. Chesebro, B. *et al. Science* **308**, 1435–1439 (2005).
6. Solfrosi, L. *et al. Science* **303**, 1514–1516 (2004).
7. Caughey, B. & Raymond, G. J. *J. Biol. Chem.* **266**, 18217–18223 (1991).

EARTH SCIENCE

Volcanic tremor wags on

A new model for volcanic tremor has a magma column, surrounded by gas bubbles, oscillating or 'wagging' back and forth. The model reconciles several observations of this characteristic signal. SEE LETTER P.522

STEPHEN R. MCNUTT

Volcanic tremor is a type of seismic signal observed at active volcanoes. On page 522 of this issue, Jellinek and Bercovici¹ present a model of tremor during eruptions that constitutes an advance in understanding this intriguing phenomenon.

The chief characteristics of volcanic tremor are its long duration (minutes to days are common) and restricted frequency range, typically 1–5 Hz, with 2–3 Hz being especially common. The seismic signals often begin gradually, and the ground vibrates continuously with few distinct phases in the signal. This means that one of the seismologist's usual primary tasks — measuring the arrival times of seismic waves — is compromised.

Tremor has been observed at more than 160 volcanoes worldwide. It is especially common just before and during eruptions, and its usefulness as a forecasting tool has long been recognized. Despite this, tremor remains poorly understood. This is no doubt due, in part, to the fact that there are several regimes or types of tremor, and different models have been developed to exploit characteristic features such as the shape and length of the magma-containing conduit, and the magma properties and gas content.

Jellinek and Bercovici's model¹ consists of a column of magma surrounded in its conduit by an annulus of gas bubbles (see Fig. 2a on page 524). The magma 'wags' back and forth in the conduit, with the gas bubbles acting as springs, providing a restoring force to keep the magma centred in the conduit.

This model has several appealing features. It is not sensitive to the exact cross-sectional shape of the conduit (a cylindrical shape has been used in the modelling for mathematical simplicity). In addition, it is faithful to recent geological observations of zones at the edges of exhumed conduits that are filled with gas and breccia (broken-up rock). The model also reconciles some features of flow, such as higher flow rates near the centre of the conduit and lower rates near the walls, where interactions take place that affect tremor production. Jellinek and Bercovici also provide a concise and up-to-date summary of various tremor attributes.

Previous explanations of tremor have focused on the geometry and size of the conduit to account for the observed frequencies. For example, a cylindrical pipe (organ pipe model; see references in ref. 1) will behave as a one-dimensional oscillator, with a frequency determined by the length and the speed of sound waves in the magma, which can be

varied by changing the concentration of bubbles. Overtones are possible, resulting in the often-observed narrow, evenly spaced spectral peaks whose frequencies may change systematically or 'glide'. Two-dimensional cracks with different lengths and widths² give rise to sets of frequencies that interact with each other, producing complex spectra and a complicated radiation pattern of the transmitted energy.

Each of these models has a fluid (the magma or water and gas) in contact with rock. A percentage of the energy in the fluid is transmitted to the rock, through which seismic waves propagate to the seismic stations. In one alternative approach³, the walls are themselves pushed apart by the magma and push back, acting as dampened springs. The resulting self-sustained oscillations are an efficient way to both generate tremor and modulate magma flow, the velocity of which changes as the walls move apart or together (the Bernoulli effect). Yet other formulations consider the effects of source and propagation factors separately; both affect the resulting tremor signal⁴.

Jellinek and Bercovici's model¹ offers alternative explanations for many of the features used to formulate these other models. An especially welcome contribution is that the main frequencies produced by wagging are in the 1–5-Hz range, exactly the dominant frequencies observed for most tremor. Importantly, these frequencies are caused by the apparent stiffness of the gas annulus (the spring) and are not related to the dimensions of the conduit. This marks a fundamental distinction between this and previous efforts.

The model also returns similar frequencies for reasonable choices of input parameters such as conduit length, shape or diameter, and is not sensitive to magma composition (andesite, dacite, rhyolite and so on). It also demonstrates that higher frequencies of tremor, up to 7 Hz or more, are produced during explosive eruptions. During eruptions, fragmentation and flow of gases occur in the annulus, causing it to be thinner and stiffer, and hence producing higher wagging frequencies. Such an increase in the frequency of tremor is observed for many eruptions.

There are several limitations to Jellinek and Bercovici's formulation¹. It may explain only one type of tremor — that during eruptions — and is unlikely to be applicable to deep tremor emanating from around 40 km depth, or tremor caused by hydrothermal boiling. And it does not explicitly address how the wagging system is coupled to the surroundings. Furthermore, the model is simplified to include mainly linear effects: nonlinear effects such as feedback may be relevant in some cases.

Nonetheless, this work¹ provides a fresh perspective on an important and long-standing problem. The basic elements of the model may also provide testable elements to provoke the next generation of field observations. ■

Stephen R. McNutt is at the Alaska Volcano Observatory, Geophysical Institute, University of Alaska Fairbanks, 903 Koyukuk Drive, Fairbanks, Alaska 99775, USA.
e-mail: steve@giseis.alaska.edu

1. Jellinek, A. M. & Bercovici, D. *Nature* **470**, 522–525 (2011).
2. Chouet, B. *Nature* **380**, 309–316 (1996).
3. Julian, B. R. J. *Geophys. Res.* **99**, 11859–11877 (1994).
4. Neubeurg, J. *Phil. Trans. R. Soc. A* **358**, 1533–1546 (2000).

BIOPHYSICS

Flipping Watson and Crick

Watson–Crick base pairs underpin the DNA double helix. Evidence of transient changes in base-pairing geometry highlights the fact that the information held in DNA's linear sequence is stored in three dimensions. [SEE ARTICLE P.498](#)

BARRY HONIG & REMO ROHS

Linear sequences of the DNA bases adenine, cytosine, guanine and thymine (A, C, G and T) define the amino-acid sequences of proteins through the genetic code. Additional codes have been sought to account for the fact that certain proteins bind to particular base sequences in DNA. It has become clear, however, that protein–DNA interactions involve subtle molecular-recognition phenomena. Further striking evidence of this is provided by Nikolova *et al.*¹ in

this issue (page 498). They have used nuclear magnetic resonance (NMR) spectroscopy to reveal intrinsic variations in the hydrogen-bonding patterns of base pairs within DNA double helices. Taken together with recent evidence^{2,3} of sequence-dependent variations in DNA shape, it now seems extremely unlikely that a simple linear code exists to determine the binding specificity of proteins for DNA.

Ten years after James Watson and Francis Crick published their model of the DNA double helix⁴, Karst Hoogsteen reported a crystal structure⁵ of a complex in which analogues of A and T formed a base pair that had a

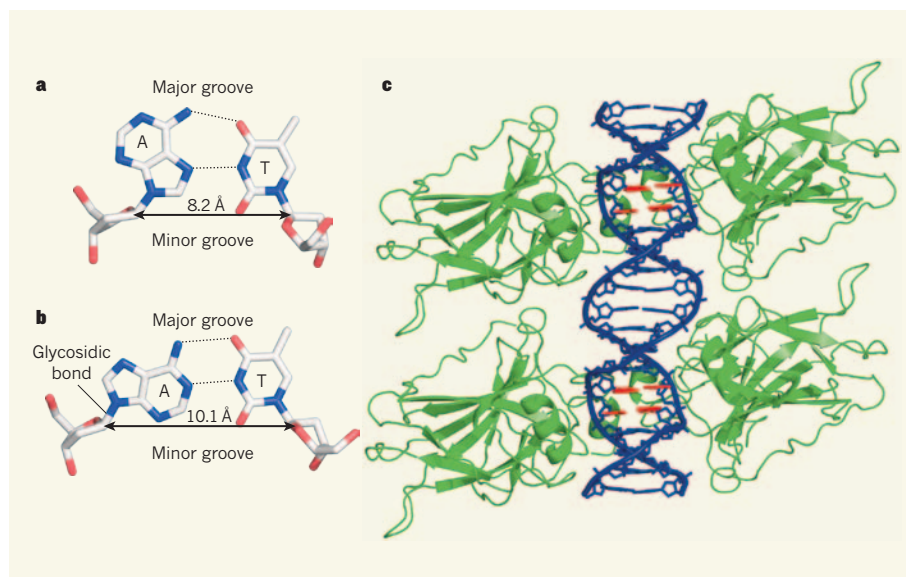


Figure 1 | Hoogsteen and Watson–Crick base pairing. Base pairs in nucleic acids can adopt different geometries. **a**, This A•T Hoogsteen base pair forms in a complex of DNA with the p53 protein. Structural data are from ref. 9. **b**, This A•T Watson–Crick base pair forms at the same position in an analogous p53–DNA complex, although the DNA has a different overall sequence. Structural data are from ref. 11. The adenine is rotated by approximately 180° about the glycosidic bond compared with the Hoogsteen base pair, so that the lower hydrogen bond involves a different chemical group in the adenine from that used in the lower hydrogen bond in **a**. The pattern of chemical groups in each groove of the DNA differs between **a** and **b**, as does the distance across the base pair. **c**, This complex⁹ consists of a tetramer of p53 DNA-binding domains (green) bound to DNA that contains Hoogsteen base pairs (red). The distortion to the double-helix structure caused by the Hoogsteen base pairs is clearly visible. Nikolova *et al.*¹ report that Hoogsteen base pairs form transiently in free DNA.

different geometry from that described by Watson and Crick (Fig. 1a, b). Similarly, an alternative base-pairing geometry can occur for G•C pairs. Hoogsteen pointed out that if the alternative hydrogen-bonding patterns were present in DNA, then the double helix would have to assume a quite different shape. Hoogsteen base pairs are, however, rarely observed.

Nikolova and colleagues' key finding¹ is that, in some DNA sequences, especially CA and TA dinucleotides, Hoogsteen base pairs exist as transient entities that are present in thermal equilibrium with standard Watson–Crick base pairs. The detection of the transient species required the use of NMR techniques that have only recently been applied to macromolecules⁶.

Why is this finding important? Hoogsteen base pairs have, after all, previously been observed in protein–DNA complexes^{7–9} (Fig. 1c). But it has not been possible to determine whether Hoogsteen base pairs are present in free DNA. Nikolova and colleagues' study¹ reveals that the ability to flip between Watson–Crick and Hoogsteen base pairing in free DNA is an intrinsic property of individual sequences. This implies that some proteins have evolved to recognize only one base-pair type, and use intermolecular interactions to shift the equilibrium between the two geometries⁹.

DNA has many features that allow its sequence-specific recognition by proteins. This recognition was originally thought to primarily involve specific hydrogen-bonding interactions between amino-acid side chains and bases. But it soon became clear that there was no identifiable one-to-one correspondence — that is, there was no simple code to be read. Part of the problem is that DNA can undergo conformational changes that distort the classical double helix. The resulting variations in the way that DNA bases are presented to proteins can thus affect the recognition mechanism.

More significantly, it has become evident that distortions in the double helix are themselves dependent on base sequence. This enables proteins to recognize DNA shape in a manner reminiscent of the way that they recognize other proteins and small ligand molecules. For example, stretches of A and T bases can narrow the minor groove of DNA (the narrower of the two grooves in the double helix), thus enhancing local negative electrostatic potentials and creating binding sites for appropriately placed, positively charged arginine amino-acid residues³.

Nikolova and colleagues' discovery¹ that DNA base pairs can so easily leave their favoured Watson–Crick conformation comes as a surprise. But viewed in another way, the phenomenon is not so different from protein side chains undergoing a conformational change so as to optimize binding with another protein. The real surprise where DNA is

concerned is that the constraints of the double helix don't preclude this possibility. The presence of Hoogsteen base pairs in detectable amounts — even in free DNA — therefore provides a notable example of the remarkable plasticity of the canonical double helix. It also implies that, if DNA bases are regarded as letters, each letter potentially has two meanings that determine both hydrogen-bonding patterns and structural variations in the double helix.

Structural biologists have long recognized that there is no second code in which certain amino acids recognize complementary DNA bases in protein–DNA interactions. Nevertheless, protein–DNA binding is still commonly thought of purely in terms of codes and sequence motifs, rather than as the binding of two large macromolecules that have complex shapes and considerable conformational flexibility¹⁰. Nikolova and colleagues' discovery reminds us that DNA offers proteins not only an enticing linear alphabet, but also a set of conformations that can be recognized in a sequence-dependent way. Understanding how the linear sequence of bases in DNA is recognized by proteins is therefore a problem that must be solved in three dimensions. This will require structural, biochemical, genomic and

computational studies on both naked double helices and protein–DNA complexes. ■

Barry Honig is at the Howard Hughes Medical Institute, Center for Computational Biology and Bioinformatics, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA. **Remo Rohs** is in the Molecular and Computational Biology Program, Department of Biological Sciences, University of Southern California, Los Angeles, California 90089, USA.

e-mails: bh6@columbia.edu; rohs@usc.edu

1. Nikolova, E. N. *et al.* *Nature* **470**, 498–502 (2011).
2. Parker, S. C. J., Hansen, L., Abaan, H. O., Tullius, T. D. & Margulies, E. H. *Science* **324**, 389–392 (2009).
3. Rohs, R. *et al.* *Nature* **461**, 1248–1253 (2009).
4. Watson, J. D. & Crick, F. H. C. *Nature* **171**, 737–738 (1953).
5. Hoogsteen, K. *Acta Crystallogr.* **16**, 907–916 (1963).
6. Palmer, A. G. & Massi, F. *Chem. Rev.* **106**, 1700–1719 (2006).
7. Aishima, J. *et al.* *Nucleic Acids Res.* **30**, 5244–5252 (2002).
8. Nair, D. T., Johnson, R. E., Prakash, S., Prakash, L. & Aggarwal, A. K. *Nature* **430**, 377–380 (2004).
9. Kitayner, M. *et al.* *Nature Struct. Mol. Biol.* **17**, 423–429 (2010).
10. Rohs, R. *et al.* *Annu. Rev. Biochem.* **79**, 233–269 (2010).
11. Chen, Y., Dey, R. & Chen, L. *Structure* **18**, 246–256 (2010).

CLIMATE CHANGE

Old droughts in New Mexico

A long climate record reveals abrupt hydrological variations during past interglacials in southwestern North America. These data set a natural benchmark for detecting human effects on regional climates. SEE LETTER P.518

JOHN WILLIAMS

Southwestern North America is a pretty dry place, and is likely to get drier this century because of anthropogenic climate warming. On page 518 of this issue¹, Fawcett *et al.* provide a climate record from deep in the past that will help in assessing the future hydrological regime for the area.

Climate models consistently project declines in winter precipitation for the southwest, in response to rising greenhouse gases as the subtropical dry zones expand polewards^{2,3}. This precipitation decline, combined with expected increases in evaporation rates and reduced snowpack, would severely strain the region's capacity to adapt to climate change. Moreover, the southwest is historically prone to droughts, with six multi-year droughts in the nineteenth and twentieth centuries, including the infamous 1930s

Dust Bowl⁴. These droughts are linked to yearly-to-decadal variations in sea surface temperatures in the tropical Pacific, enhanced by local soil-moisture feedbacks⁵. Further understanding of the mechanisms of hydrological variability, together with efforts to limit societal vulnerability to climate change, are priorities in global-change research⁴.

Palaeoclimatic studies have made an essential, if not particularly reassuring, contribution to this effort, by providing insight into the natural behaviour of hydrological systems and a longer-term context for historical and projected changes. Tree-ring records offer sobering evidence of widespread and decades-long 'megadroughts' in the western United States over the past several millennia that dwarfed recorded historical droughts⁶. Records spanning the past 11,000 years (the Holocene interglacial) demonstrate long-term shifts in southwestern monsoon



Figure 1 | An aerial view of Valles Caldera, site of the lake sediments from which Fawcett *et al.*¹ extracted their climate record.

intensity⁷. Fawcett *et al.*¹ now present a beautifully detailed record of climatic variability between 370,000 and 550,000 years ago from lake sediments at Valles Caldera, New Mexico (Fig. 1). This time interval spans two earlier interglacial periods, known as Marine Isotope Stage (MIS) 11 and 13.

Among palaeoclimatic aficionados, MIS 11 is of great interest, because Earth's orbital configuration, and hence the insolation regime (amount of solar radiation reaching Earth), closely resembles that during the Holocene⁸. Usually, interglacial periods are short, lasting half of one precessional cycle (about 10,000 years), and end when summer cooling in the Northern Hemisphere triggers a new round of continental glaciations. However, during MIS 11, Earth's orbit was nearly circular for an unusually long time, dampening the effect of precessional variations on seasonal insolation and causing MIS 11 to last 50,000 years. (MIS 13 is comparably long and hence also of interest.) The similar orbital configuration today raises the fascinating possibility that, even without rising greenhouse gases, the current interglacial might have another 40,000 years to go⁹. Thus, palaeoclimatic records from MIS 11 offer unique insights into the potential trajectory for the contemporary climate system, in the absence of humanity's effects.

Unfortunately, most MIS 11 records are from Antarctic ice cores or marine sediments; terrestrial records are scarce. The Valles Caldera record is remarkable for the wealth of palaeoclimatic proxy data collected by Fawcett and colleagues¹. The proxies include both new geochemical indicators of past temperatures and tried-and-true proxies for terrestrial vegetation (pollen and $\delta^{13}\text{C}$ [$^{13}\text{C}/^{12}\text{C}$]), lake productivity (silica/titanium ratios) and

lake hydrology (total organic carbon and calcium concentrations). The age model is well constrained by a basal, radiometrically dated ash layer and by palaeomagnetic reversals of known age, and the temporal resolution is fine enough to resolve millennial-scale climate variations.

The Valles Caldera record reveals three warm and two cool stages within MIS 11, with a 2 °C difference between stages. Warm peaks probably correspond to peaks in summer insolation.

The first warm stage (MIS 11e) is the counterpart to the past 11,000 years, and its trajectory is documented in exquisite detail. MIS 11e began with a rapid 8 °C warming, increased abundances of vegetation such as oak and juniper that prefer warm conditions, and increased lake productivity. A few thousand years later, abundances of grasses with the C_4 photosynthetic system increased, suggesting warm and at least seasonally wet conditions. Simultaneously, however, the Valles Caldera lake became hydrologically closed (there were no outflowing streams), a signal of negative water balance. This somewhat paradoxical result can be explained by differing seasonal precipitation signals, with the C_4 grasses responding to continued summer precipitation and lake hydrology responding to reduced winter precipitation. Summer and winter precipitation also diverged during the early Holocene, probably due to the effects of insolation on southwestern monsoon intensity¹⁰.

The next phase within MIS 11e is the most striking: abrupt shifts in $\delta^{13}\text{C}$ and lowered total organic carbon suggest a rapid collapse of C_4 grasses and oxidation of lake sediments; mud cracks in the sediments show that the lake dried out. This arid episode lasted several

thousand years, and apparently began and ended abruptly. At least two other multi-millennial dry periods occurred during MIS 11, and higher-frequency variations in total organic carbon and $\delta^{13}\text{C}$ hint at sub-millennial hydrological variability during the warm stages of MIS 11.

Mud cracks and low total organic carbon also characterize much of MIS 13, implying that this interglacial was similarly marked by recurring episodes of high temperatures and aridity. Interestingly, MIS 13 seems to have been warmer and drier than MIS 11, despite lower concentrations of atmospheric carbon dioxide and methane during MIS 13 (refs 11, 12). This suggests that the larger variations in insolation during MIS 13 strongly regulated southwestern aridity.

What lessons can we draw from the Valles Caldera record? First, in southwestern North America, hydrological variability seems to be the rule rather than the exception during interglacial periods. Second, on timescales of 10^3 to 10^4 years, orbital precession strongly influences southwestern monsoon intensity and seasonal precipitation. Third, insofar as MIS 11 is a good analogue for the Holocene, we should now be at a time roughly equivalent to the transition between warm MIS 11e and cool MIS 11d. If so, southwestern climates might naturally be trending towards a somewhat cooler and wetter stage — except that other factors (us) are affecting climate. Knowing the likely natural trends thus helps us to diagnose the causes of twenty-first-century hydrological trends. And perhaps, just perhaps, these natural trends will partially mitigate the projected drying in the southwest. ■

John (Jack) Williams is in the Department of Geography and Center for Climatic Research, University of Wisconsin-Madison, Madison, Wisconsin 53706-1695, USA.
e-mail: jww@geography.wisc.edu

1. Fawcett, P. J. *et al.* *Nature* **470**, 518–521 (2011).
2. Intergovernmental Panel on Climate Change *Climate Change 2007: The Physical Science Basis* (Cambridge Univ. Press, 2007).
3. Seager, R. & Vecchi, G. A. *Proc. Natl Acad. Sci. USA* **107**, 21277–21282 (2010).
4. Cook, E. R. *et al.* in *Abrupt Climate Change. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research* (eds Clark, P. U. *et al.*) 143–257 (US Geol. Surv., 2008).
5. Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D. & Bacmeister, J. T. *Science* **303**, 1855–1859 (2004).
6. Cook, E. R., Seager, R., Cane, M. A. & Stahle, D. W. *Earth Sci. Rev.* **81**, 93–134 (2007).
7. Thompson, R. S., Whitlock, C., Bartlein, P. J., Harrison, S. P. & Spaulding, W. G. in *Global Climates Since the Last Glacial Maximum* (eds Wright, H. E. Jr *et al.*) 468–513 (Univ. Minnesota Press, 1993).
8. Loutre, M. F. & Berger, A. *Global Planet. Change* **36**, 209–217 (2003).
9. Berger, A. & Loutre, M. F. *Science* **297**, 1287–1288 (2002).
10. Harrison, S. P. *et al.* *Clim. Dyn.* **20**, 663–688 (2003).
11. Siegenthaler, U. *et al.* *Science* **310**, 1313–1317 (2005).
12. Spahni, R. *et al.* *Science* **310**, 1317–1321 (2005).

A turbulent matter

A long-standing controversy about whether the motions within a typical astrophysical disk of gas are stable or unstable has resurfaced. The answer has profound significance for our understanding of how stars and planets form.

STEVEN A. BALBUS

Writing in *Physical Review Letters*, Paoletti and Lathrop¹ describe the breakdown of smooth flow into turbulence in a laboratory investigation. What makes this so noteworthy is that the flow in question is similar to what is generically expected to be present in astrophysical disks². Moreover, the authors reach precisely the opposite conclusion to what had been considered the state-of-the-art laboratory investigation³, which found no such breakdown into turbulence and implied that, in the absence of an external agent such as a magnetic field, the rotation pattern of astrophysical disks should be stable.

Understanding how a fluid with internal motion makes a transition from smooth (laminar) flow to turbulence remains a stubborn, long-standing theoretical challenge to the hydrodynamics community⁴. A particularly relevant astrophysical application of this problem is the study of star formation, because stars similar to the Sun are thought to pass the earliest stages of their lives forming at the centres of gaseous 'protostellar' disks⁵. The question is whether the gas, which orbits in nearly circular motion (much like the planets orbiting the Sun), is a turbulent or laminar fluid. Depending on the answer to this question, the astrophysical consequences may be very different. The efficiency with which a forming star is able to accrete gas from a surrounding disk is much greater if the gas is turbulent, for example. Furthermore, planet formation in the disk may be aided by the gathering of dust grains into transient whirlpool structures⁶.

The formal problem of the stability of rotating flow was first addressed by Lord Rayleigh in the late nineteenth century⁷. Rayleigh found that if the rotational velocity of a fluid decreases more rapidly with radius than the reciprocal of the distance from the axis of rotation, such a system is unstable to infinitesimal perturbations. Astrophysical disks, by this criterion, should be stable. But Rayleigh's analysis was restricted to vanishingly small disturbances, and the geometrical shape of the perturbations was in the form of rings with cylindrical symmetry. It is still not known what types of flow that are formally stable by this Rayleigh criterion might still be unstable to more general forms of disturbance; it is known, however, that some types of Rayleigh-stable

flow certainly can be destabilized^{4,8}. The issue of interest is whether the rotation of an astrophysical gas disk about a central mass falls into this unstable category.

This problem can be investigated in the laboratory by studying what is known as Couette flow. In a Couette apparatus, water is confined to flow in the space between two coaxial cylinders. There should be no motion along the central axis, only rotational flow about the axis. The cylinders rotate independently of one another, so that small frictional viscous forces near the cylindrical walls will set up a hydrodynamical flow in which the rotational velocity depends on the distance from the rotation axis. By choosing the rotational velocities of the rotating cylinders appropriately, a small section of an astrophysical disk can be mimicked in the laboratory. In such a disk, the flow velocity is inversely proportional to the square root of the distance from the centre, a pattern known as Keplerian flow. The question to be answered is whether Keplerian flow, formally stable by the Rayleigh criterion, actually breaks down into turbulence.

It is this question that Paoletti and Lathrop¹ have sought to address. When a Couette flow becomes turbulent, one of the consequences is a greatly enhanced outward flux of angular momentum, which is imparted to the outer cylinder in the form of a torque. In their experiment, the authors measure this torque directly. An earlier investigation⁹ had claimed to detect this torque, but the new experiment¹ was conducted under conditions in which (undesirable) viscous effects were more effectively minimized.

Close on the heels of Paoletti and Lathrop's claim, however, comes a report by Schartman *et al.*¹⁰ on a related experiment. These investigators found no transition to turbulence for Keplerian flow with the same controlled level of viscosity. This null result was first reported³ in 2006, and the most recent paper maintains its original conclusion that there is no evidence of a turbulent breakdown of Keplerian-like laminar flow for very small values of the viscosity.

Although the two experiments^{1,10} are in principle investigating the same type of rotational flow profile, the measured quantities and the Couette apparatuses themselves are different. Paoletti and Lathrop¹ work with a container in which the end caps at the top and bottom of the cylinders rotate as solid

disks. The fluid immediately adjacent to an end cap also rotates uniformly and therefore must make a sharp transition, within a narrow region, to join onto the shearing rotation in the bulk of the flow. The transition causes additional stresses to be present throughout the fluid, which in turn drive vertical circulation patterns (technically known as Ekman flow) that would not be present in astrophysical disks. To avoid this unwanted Ekman flow, the authors¹ work with long cylinders, so that near the midplane of the apparatus these effects are minimized.

Schartman *et al.*¹⁰ use instead a Couette system with split end caps comprised of two sliding annuli, so that four velocities can be adjusted: those of the two confining cylinders, and the rotation rates of the annuli. The Ekman circulation may thus be directly controlled. The experiment¹⁰ is also equipped to measure the internal velocity of the fluid directly. This ability serves to verify that the correct rotation profile has been achieved, as well as to detect the characteristic fluctuations in velocity that accompany a breakdown into turbulence. Paoletti and Lathrop's experiment¹, by contrast, is designed to measure only the enhanced torque, not the velocity field itself.

The two approaches have led their respective investigative teams to exactly opposite conclusions. Paoletti and Lathrop argue that their results show that astrophysical disks would be unstable to large-amplitude disturbances (as opposed to the infinitesimal perturbations assumed in Rayleigh's mathematical analysis) and become turbulent. By contrast, Schartman *et al.* maintain that there are no significant dynamical instabilities in which the Keplerian rotation field of the Couette flow — or, presumably, of an astrophysical disk — is the driving source of energy for the onset of turbulence.

By way of support, Paoletti and Lathrop can point to a recent fluid experiment by van Gils *et al.*¹¹ that finds the same quantitative relationship between the rotation parameters of the cylinders and the ensuing turbulent torque. It must be noted, however, that in the study by van Gils and colleagues, the unstable profiles are not near the Keplerian regime. On the other hand, the null result of Schartman *et al.* is itself supported by direct numerical simulations^{12,13} showing stability of the flows in question. Here, the caveat is that the simulations do not yet have viscous effects controlled at the same level that the laboratory experiments can now achieve.

Because of its central importance to astrophysics, the possibility that disks may be turbulent for purely hydrodynamical reasons will probably excite another round of intense investigative activity, both in the laboratory and on the computer. For the time being, however, we must wait a little longer for a laboratory consensus on whether Keplerian disks are, after all, intrinsically unstable, or whether, as

is currently suspected by most accretion-disk theorists, magnetic effects have an essential role in the destabilization process. ■

Steven A. Balbus is at the *Laboratoire de Radioastronomie, École Normale Supérieure, 75231 Paris Cedex 05, France.*
e-mail: steven.balbus@lra.ens.fr

1. Paoletti, M. S. & Lathrop, D. P. *Phys. Rev. Lett.* **106**, 024501 (2011).
2. Frank, J., King, A. & Raine, D. *Accretion Power in Astrophysics* 3rd edn (Cambridge Univ. Press, 2002).
3. Ji, H., Burin, M., Schartman, E. & Goodman, J. *Nature* **444**, 343–346 (2006).
4. Drazin, P. G. & Reid, W. H. *Hydrodynamic Stability* (Cambridge Univ. Press, 1981).

5. Hartman, L. *Accretion Processes in Star Formation* (Cambridge Univ. Press, 2000).
6. Johansen, A. *et al. Nature* **448**, 1022–1025 (2007).
7. Lord Rayleigh *Proc. R. Soc. Lond. A* **93**, 148–154 (1917).
8. Coles, D. J. *Fluid Mech.* **21**, 385–425 (1965).
9. Richard, D. *Instabilités Hydrodynamiques dans les Écoulements en Rotation Différentielle*. PhD thesis, Univ. Paris 7 (2001).
10. Schartman, E., Ji, H., Burin, M. J. & Goodman, J. *Astron. Astrophys.* (submitted); preprint available at <http://arxiv.org/abs/1102.3725> (2011).
11. van Gils, D. P. M., Huisman, S. G., Bruggert, G.-W., Sun, C. & Lohse, D. *Phys. Rev. Lett.* **106**, 024502 (2011).
12. Hawley, J. F., Balbus, S. A. & Winters, W. F. *Astrophys. J.* **518**, 394–404 (1999).
13. Lesur, G. & Longaretti, P.-Y. *Astron. Astrophys.* **444**, 25–44 (2005).

CANCER

When catastrophe strikes a cell

In 2–3% of cancers, a single genetic event may have led to hundreds of genomic rearrangements confined to just one or a few chromosomes. This finding challenges the conventional view of how mutations accumulate in oncogenesis.

JOSE M. C. TUBIO & XAVIER ESTIVILL

How do the mutations that lead to cancer come about? The traditional view¹ is that a gradual process involving continual acquisition of heritable genetic changes by cells causes cancer. There is, however, an alternative view that single catastrophic events can lead to multiple mutations. In a paper published in *Cell*, Stephens *et al.*² provide evidence for the concept of catastrophism in cancer.

Our understanding of oncogenesis has benefited greatly from next-generation

sequencing technology. Stephens and colleagues combined next-generation sequencing² and single nucleotide polymorphism (SNP) array data^{3,4} to analyse the patterns of somatic (non-germline) genomic rearrangements in tumours. Intriguingly, they found that in some cases the changes consist of tens to hundreds of rearrangements, confined to one or a few chromosomes. The authors² coin the term chromothripsis (from ‘chromo’, for chromosome; and ‘thripsis’, for breaking into small pieces) to designate this phenomenon². Although they observed evidence of a high incidence of chromothripsis in bone tumours, it

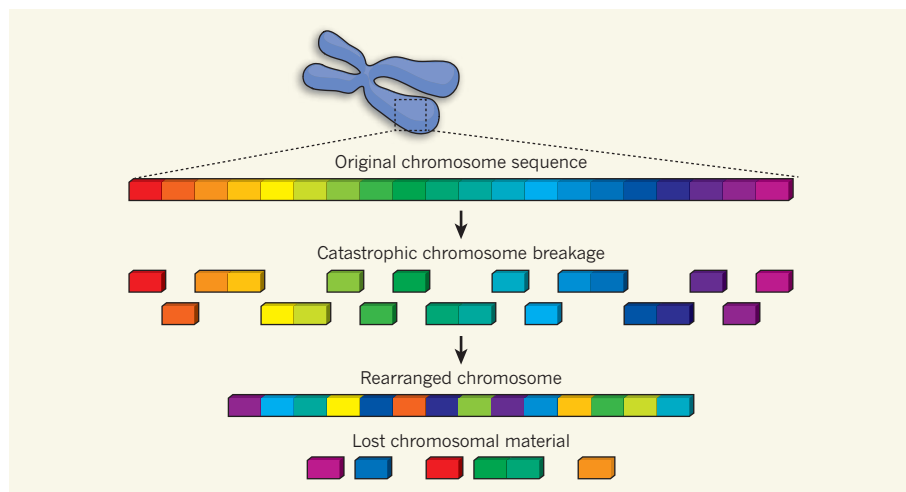


Figure 1 | Chromothripsis. Stephens *et al.*² report that, when a catastrophic event causes DNA fragmentation, its subsequent repair leads to chromosomal rearrangements, as well as the loss of some sequences. They call this phenomenon chromothripsis. (Adapted from ref. 2.)

seems to occur in at least 2–3% of all cancers.

Rearrangements are common in cancer, so how can Stephens *et al.* tell whether such mutations were caused by a single event? They argue that the final configuration of the rearrangements they observed could be explained only by a single catastrophic episode, rather than by a series of independent events. Rearrangements due to chromothripsis are usually restricted to a few chromosomes, within which breakpoints show a non-random distribution.

In a patient with leukaemia, for example, the authors found seven rearrangements clustered within a region spanning just 30,000 nucleotides. Stephens and co-workers propose that when a chromosome or chromosomal region shatters into tens or hundreds of fragments, the DNA-repair machinery reassembles some of the fragments incorrectly, leading to rearrangements (Fig. 1).

Another of the authors' observations also favours chromothripsis as the initiating event — in the final configuration, the copy number varies along each affected chromosome arm, usually alternating between just one and two copies. By contrast, statistical simulations showed that, if the rearrangements had been acquired gradually, tandem duplications should have increased the copy number of the associated genomic segments several-fold.

The potential implications of chromothripsis as a cause and/or mediator of cancer are evident. As Stephens *et al.* point out, the generation of so many rearrangements in a single genomic crisis makes it likely that more than one cancer-causing lesion would occur. Indeed, in their samples they describe several cancer-related genes that are affected by the rearrangements. In one patient, for instance, they found that a single catastrophic event resulted in the disruption of three tumour-suppressor genes — *CDKN2A*, *FBXW7* and *WRN*. In addition, on the basis of several cases, it seems that chromothripsis rearrangements can produce potentially oncogenic fusion genes, by juxtaposing coding regions of two different genes.

Stephens *et al.* clearly establish the hallmarks of chromothripsis, but the mechanism and the cause of such DNA fragmentation remain open for discussion. Because chromothripsis rearrangements are strictly limited to one chromosome or chromosomal region, the authors posit that catastrophe strikes when chromosomes condense for mitotic cell division. They assume that if the damage occurred during the interphase stage of the cell cycle — when the DNA structure is more relaxed — it is unlikely that it would lead to clusters of breaks within such well-circumscribed genomic regions. An alternative explanation for spatially restricted chromothripsis could simply be that several damaged chromosomes would decrease the probability of the cell's survival (Fig. 2).

As for the cause of chromothripsis, Stephens

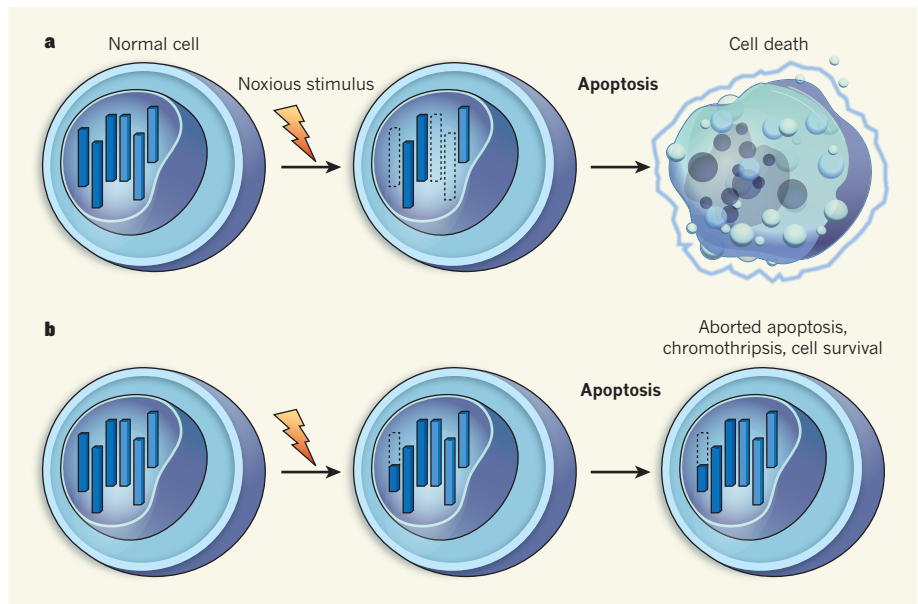


Figure 2 | Severity of damage matters. On exposure to a noxious stimulus, programmed cell death (apoptosis) probably commences, triggering DNA fragmentation (dotted regions). **a**, Consequently, most cells die. **b**, If the damage is not too severe, however, abortion of apoptosis could occur, and incorrect repair of DNA fragments could lead to the chromosomal rearrangements characteristic of chromothripsis. The surviving cells could eventually become cancerous.

et al. suggest two possibilities. For one, ionizing radiation could induce chromosome breaks. Depending on how the radiation affects chromosome structure, breaks could occur in a short or a long stretch of the chromosome. Another possibility is dysfunction of telomeres — cap-like nucleoproteins at the tips of chromosomes. Telomere dysfunction is known⁵ to promote chromosomal abnormalities that often typify cancer cells, including end-to-end chromosome fusion, anaphase bridges, aneuploidy and polyploidy. This possibility is appealing given that, in most cases, chromothripsis involves regions that extend to telomeres.

We propose a third possibility: aborted programmed cell death (apoptosis) might cause chromothripsis (Fig. 2). There is evidence⁶ that the abortion of apoptosis in its initial stages may lead to chromosomal rearrangements in cancer. Accordingly, it could be that noxious stress stimuli (such as radiation, nutrient deprivation, infection or oxygen shortage) induce apoptosis in a cell population, initiating higher-order fragmentation of chromatin (DNA–protein complexes). For most of the population, the outcome would be relentless cell death. But somehow one or a small subset of cells may not complete apoptosis and so may survive. The surviving cells would then need to repair the cleaved DNA, but some might do so incorrectly, leading to rearrangements.

This alternative hypothesis could explain the complex chromosomal configuration that characterizes chromothripsis². As for the significant non-random distribution of breakpoints along chromosomes, this could be due to targeted cleavage and/or cleavage

in the most exposed DNA regions^{7–9}.

What could drive the abortion of apoptosis? Many members of the γ -herpesvirus family can cause cancer, possibly by inhibiting apoptosis¹⁰. Could such viruses also be involved in chromothripsis? With the large amount of unbiased data that can be generated through sequencing in cancer projects, testing this possibility should be feasible.

The authors' results² indicate that chromothripsis can have cancer-promoting genomic consequences, but could equally be interpreted as implying that chromothripsis is a consequence of the initial stages of tumorigenesis. These observations are probably a snapshot of the cellular changes that occur in some cases of oncogenesis. ■

Jose M. C. Tubio and Xavier Estivill are in the Genes and Disease Programme, Center for Genomic Regulation (CRG) and Pompeu Fabra University, Barcelona, Catalonia 08003, Spain.
e-mail: xavier.estivill@crg.cat

- Stratton, M. R., Campbell, P. J. & Futreal, P. A. *Nature* **458**, 719–724 (2009).
- Stephens, P. J. *et al.* *Cell* **144**, 27–40 (2011).
- Bignell, G. R. *et al.* *Nature* **463**, 893–898 (2010).
- Beroukhi, R. *et al.* *Nature* **463**, 899–905 (2010).
- Titen, S. W. A. & Golik, K. G. *Genetics* **180**, 1821–1832 (2008).
- Stanulla, M., Chhailiyil, P., Wang, J., Jani-Sait, S. N. & Aplan, P. D. *Hum. Mol. Genet.* **10**, 2481–2491 (2001).
- Czene, S., Testa, E., Nygren, J., Belyaev, I. & Harms-Ringdahl, M. *Biochem. Biophys. Res. Commun.* **294**, 872–878 (2002).
- Cremer, T. & Cremer, C. *Nature Rev. Genet.* **2**, 292–301 (2001).
- Liu, Q. Y. *et al.* *Cell Death Differ.* **10**, 278–289 (2003).
- Lagunoff, M. & Carroll, P. A. *Int. Rev. Immunol.* **22**, 373–399 (2003).

Climate change and evolutionary adaptation

Ary A. Hoffmann¹ & Carla M. Sgrò²

Evolutionary adaptation can be rapid and potentially help species counter stressful conditions or realize ecological opportunities arising from climate change. The challenges are to understand when evolution will occur and to identify potential evolutionary winners as well as losers, such as species lacking adaptive capacity living near physiological limits. Evolutionary processes also need to be incorporated into management programmes designed to minimize biodiversity loss under rapid climate change. These challenges can be met through realistic models of evolutionary change linked to experimental data across a range of taxa.

Natural populations are responding to global climate change by shifting their geographical distribution and timing of growth and reproduction, and these changes are, in turn, altering the composition of communities and the nature of species interactions¹. However, the responses of many populations are likely to be inadequate to counter the speed and magnitude of climate change, leaving groups such as lizards vulnerable to decline and extinction². Extinction can be avoided if populations move to favourable habitats, organisms successfully overcome stressful conditions via plastic changes, or populations undergo evolutionary adaptation³.

Recent studies have highlighted that evolutionary change can be rapid in a number of taxa⁴, including in species that have invaded new areas⁵ and in native species responding to biotic invasions⁶. This indicates that evolutionary adaptation could be an important way for natural populations to counter rapid climate change, and that predicted colonization patterns and distribution shifts are markedly affected by the inclusion of evolution^{7,8}. Evolutionary adaptation might be the only way that threatened species can persist if they are unable to disperse naturally or through human-mediated translocation to climatically suitable habitats. This process might also be essential for the ongoing health of keystone species facing threats arising from climate change, as in the case of dominant conifers being attacked by bark beetle populations benefiting from warming conditions⁹. Adaptive changes are likely to influence the ability of species to take advantage of potentially favourable conditions arising from climate change, including the effects of CO₂ enrichment on growth rate¹⁰ and the extension of favourable seasonal conditions¹¹. However, with few exceptions, the importance of evolution tends to be ignored both in broader discussions about the effects of climate change on biodiversity and in models for predicting species responses to climate change.

Rapid climate change is likely to produce a range of new selection pressures on populations. Increasing periods of thermal stress and drought will produce directional selection for resistance, particularly in species close to physiological limits such as warm-adapted porcelain crabs¹² and tropical lizards¹³ existing close to upper thermal thresholds that are likely to be exceeded in the next few decades. Warmer conditions will select for earlier emergence from winter diapause¹¹ and for dispersal to new environments¹⁴ as species take advantage of conditions that have become newly favourable. Not all selection pressures arising from climate change will be directional; stabilizing selection is likely to occur on hatching/breeding times in birds colonizing new areas where food supply and breeding times coincide¹⁵.

Climate change is occurring at a time when natural environments are becoming increasingly fragmented through habitat destruction, and when species are being moved inadvertently or deliberately around the globe at ever faster rates^{16,17}. This means that the effects of climate change are occurring at a time when many populations are already under pressure from invading species and disturbances. Fragmentation and invasions also affect evolutionary processes by changing the way genes move around landscapes and by introducing novel genotypes into populations through hybridization.

This review considers the likelihood that evolutionary changes within species can contribute to species adapting to global climate change. Evidence for and against recent evolutionary adaptation is briefly discussed, along with the relevant evolutionary approaches and models that predict future evolutionary potential. Impacts of evolution on predicted changes in species distributions are discussed, as well as management practices that might facilitate evolutionary adaptation essential for long-term species persistence.

Plastic versus genetic change in time and space

The evolutionary potential of populations can be assessed in several ways (Table 1). Longitudinal studies of single populations help determine whether changes in traits have evolved (are genetic) or instead have occurred through plasticity (determined by the environment) (Table 1). Quantitative genetic models are often used to isolate genetic effects but these have been applied incorrectly in several longitudinal studies, resulting in phenotypic changes due to environmental effects being interpreted as genetic changes¹⁸.

Where genetic and plastic contributions have been separated successfully, the latter often seem to be more important. This pattern has been noted for changes in bird population breeding dates¹⁹ and for an increase in adult body size in marmots driven by early emergence from hibernation rather than genetic changes²⁰. Also, a decrease in body size in a Soay sheep population was caused by environmental effects that decreased the growth rate of lambs, despite the presence of heritable variation for size²¹ and selection favouring larger size.

On the other hand, there is abundant evidence for genetic adaptation to climatic conditions varying in space rather than time, particularly in cases where transplant/common garden experiments have been carried out along environmental gradients (Table 1). Transplants have been widely used in plants to demonstrate genetic adaptation to climatic changes involving altered aridity and thermal conditions^{22,23}. Common

¹Departments of Zoology and Genetics, Bio21 Institute, The University of Melbourne, Victoria 3010, Australia. ²School of Biological Sciences and Australian Centre for Biodiversity, Monash University, Victoria 3800, Australia.

Table 1 | Approaches for predicting and describing evolutionary responses to recent climate change in natural populations

Approach	Outcome	Limitation	Example references
Longitudinal studies testing for genetic changes in populations	Where relationships among individuals are available or genotypes can be repeatedly sampled and tested in a common environment, changes can be partitioned into genetic and environmental components.	Requires longitudinal and relatedness data or common environment comparisons, only possible for some populations.	21,85
Spatial studies across climatic gradients involving transplants or common garden experiments	Indicates the relative importance of plastic and genetic contributions to variation along a gradient. Can be used to identify patterns of selection.	May not indicate the speed at which adaptation occurs, patterns of local adaptation will depend on gene flow, transplants only possible for plants/relatively sessile organisms unless these can be confined.	24,43
Standing quantitative genetic variation estimates within populations	Heritable variation reflects evolutionary potential.	Only possible for subset of species, evolutionary costs of adaptation are not commonly measured.	26,86
Quantitative genetic variation estimated through selection experiments	Species/populations with genetic variation respond to selection, others do not.	Can be slow, requires multiple generation selection experiments, only possible for a subset of species, intense artificial selection and laboratory adaptation may mean results are not reflective of processes in natural populations.	28,37
Experimental evolution in simulated environments	Populations exposed to conditions relevant to climate change indicate potential for selection, may include biotic interactions.	Can be slow across multiple generations, may reflect genetic variation in direction of selection, but target of selection can be unclear.	87,88
Evidence of loss of function of candidate gene/protein	Where function has been lost, may indicate fundamental limits and loss of evolutionary potential.	Depends on understanding the biochemical/molecular basis of key ecological traits.	67
Genetic variation in candidate genes for traits pointing to potential for evolution	Where candidate genes control a substantial amount of the phenotypic variance in traits associated with climate adaptation, these might be directly assessed to indicate evolutionary potential.	Requires important candidates to be identified, applicable where phenotypic variation is controlled by major genes or their expression.	89,90

garden experiments have been applied to animal as well as plant populations to assess the relative importance of plastic and genetic effects. For instance, in common garden experiments with *Drosophila melanogaster*, environmental rearing conditions—which have a large impact on thermal resistance²⁴—contribute around one-and-a-half times more than genetic factors to clinal variation in resistance²⁵. Plastic responses themselves can evolve and contribute to climatic adaptation, as in the case of cabbage white butterfly populations, which differ in their plastic responses to temperature²⁶.

In an attempt to demonstrate evolutionary responses to climate change, genetic differences in space have been compared over time in a few cases (that is, combining the first two approaches listed in Table 1). These have provided the strongest evidence for evolutionary responses in traits related to the timing of activity or reproduction, including shifts in diapause patterns in mosquitoes¹¹, flowering time in plants²⁷ and migration patterns in birds²⁸. Changes in spatial patterns can only be established when experimental investigations are repeated across years or when stored material such as seed is available²⁷, and when the direction of selection imposed through climate change is clear²⁸.

Indirect evidence can point to evolution as a probable explanation for longitudinal changes in trait patterns across gradients. For instance, several species of passerine birds from the east coast of Australia show a latitudinal cline for body size (smaller in tropical localities); this cline has undergone a shift over the last 100 years equivalent to 7.8° in latitude, and this is probably due to evolution because nutrition has not changed²⁹. However, definitive evidence of longitudinal changes requires genetic effects to be clearly separated from environmental effects¹⁹. This may require repeated common garden experiments, whereas morphological or phenological data are normally only collected at the phenotypic level.

It has been suggested that climate change adaptation is more likely to involve altered timing of life history traits rather than evolved physiological responses¹¹, but this assertion seems premature. In pitcher-plant mosquitoes in which environmental diapause triggers have evolved, selection for phenological changes is strong³⁰ but selection for tolerance is weak³¹, which may account for a lack of physiological evolution. Evolved tolerance changes will be more difficult to demonstrate than evolved timing changes because of the low repeatability of physiological assays. Nevertheless, evolved shifts in physiological limits will be critical in avoiding extinction if populations living close to their physiological

limits are to persist under increasingly stressful conditions^{2,12}. This might be achieved through evolving increased resistance within a life-cycle stage or through evolving a way of evading stressful conditions by entering an inactive stage such as diapause resistant to stressful conditions³².

The absence of genetic adaptation to recent climate change now documented in some populations raises questions about the conditions under which evolution might be expected. Which factors might constrain or promote evolutionary responses? When they occur, will these responses be sufficient to keep up with rapid climate change?

Predicting evolution within populations

The evolutionary potential of populations to selection pressures associated with climate change can be measured through family studies and selection experiments in laboratory and field populations (Table 1). A simple model for interpreting the results of these experiments involves the breeder's equation, which predicts the size of the selection response given a certain selection pressure and trait heritability (Box 1 and Table 2). This equation applies across a single generation, although it has also been applied across multiple generations on the assumption that heritability and selection intensity remain constant. It can help highlight cases where a selection response is not possible owing to a low level of heritable variation³³ and/or rapid environmental changes and strong selection pressures².

A low genetic variance is usually regarded as unlikely to account for a poor selection response because most traits have quite high levels of genetic variation³⁴. Widespread species often have abundant genetic variation for numerous traits likely to be involved in climatic adaptation including seasonal timing^{11,35}, morphological variation affecting thermal responses³⁶, and resistance to stressful climatic conditions both in insects³⁷ and in plants³⁸. Genetic variation allows for local adaptation to climatic conditions in fitness-related traits²² including traits potentially related to physiological limits as well as phenological timing^{23,39}. On the other hand, some traits do lack genetic variation; the heritability of morphological traits and the timing of breeding can be very low in field populations^{40,41}. Moreover, the heritability of desiccation resistance and cold resistance is very low in sensitive *Drosophila* species restricted to the wet tropics, in sharp contrast to their more climatically widespread (and more resistant) relatives⁴². A low heritability in one trait might reflect a lack of genetic variation generally due to low population size, but this is not the case in sensitive *Drosophila* species⁴².

BOX 1

Population persistence models

Several models have been used to predict the response of traits to selection under climate change and to determine whether populations can successfully evolve through key traits to keep up with the rate of climate change. In the simplest case, the response to selection (R) (change in mean of a trait) is given by the breeder's equation, $R = h^2 S$, where h^2 is heritability and S is selection differential, expressed as difference in selected mean versus mean of base population. This equation can be used to predict the response of a population to climate change across a generation and to test whether the response is adequate, assuming that one trait has overriding importance in population survival².

When multiple traits are considered, the multivariate equivalent of the breeder's equation is used, where $\Delta \mathbf{z} = \mathbf{G}\boldsymbol{\beta}$, relating a vector of changes in trait means ($\Delta \mathbf{z}$) to the genetic variance–covariance (\mathbf{G}) matrix and to the vector of linear selection gradients ($\boldsymbol{\beta}$). There is ongoing debate about how long the genetic variance–covariance matrix stays the same under multiple generations of selection. Where multiple traits influence fitness, the response to selection can depend on how interactions among traits influence fitness. Several ways of assessing the predicted changes in traits can then be applied⁴³, such as the change in trait means, $\Delta \bar{\mathbf{z}}$, through $\Delta \bar{\mathbf{z}} = \text{covA}[\mathbf{W}, \mathbf{z}]$, where covA is the additive genetic covariance between all pairs of traits, w is individual relative fitness and \mathbf{z} is the vector of traits.

Over multiple generations, the ability of a population to counter the effects of climate change will depend on population size. The critical rate of environmental change where population growth rate is maintained to be greater than 0 (k_c) is given by

$$k_c = \frac{V_A}{\sqrt{V_W}} \sqrt{2r_{\max} - \frac{V_P}{V_W} - \frac{1}{2N_E}}$$

where N_E is the effective population size, V_W the width of the selection function, V_P the phenotypic variance, V_A the additive genetic variance, r_{\max} the maximum growth without selection⁴⁷. Other models incorporating demographic factors have also been developed⁵¹.

Alternative formulations for critical environmental rate incorporating plasticity have also been developed, such as

$$k_c = \sqrt{\frac{2r_{\max}V_W}{T}} \frac{h^2 V_P}{|B - b|}$$

where T is the generation time, B measures how environmental changes influence the optimal phenotype, and b measures plasticity⁵⁰. A cost to plasticity can be included in formulations of these models.

A different class of models considers the ability of mutations to rescue asexual populations from extinction after climate change. Rescue only occurs when the fraction of rescue mutations (ϕ) as a proportion of the overall genome mutation rate (U) is defined as $\phi > -[r_0/(r_0 - r_1)][\ln(1 - P)]/(2N_0U)$, where r_0 is the growth rate due to the stress, r_1 is the growth rate of the rescue mutations, N_0 is the initial carrying capacity and P is the probability of spread of the beneficial mutation¹⁰.

Evolutionary changes take place in a multivariate space, where both genetic interactions between traits and the direction of selection across multiple traits dictate the potential for selection responses (Table 2). These factors can markedly slow rates of evolutionary adaptation³⁴, as demonstrated in some plants responding to rapid climate change⁴³. Trait interactions may restrict shifts in breeding time in birds⁴⁴ as well as thermal adaptation in insects because genes increasing resistance to thermal stresses can have relatively lower fitness under favourable conditions in the laboratory and in nature^{45,46}. Genetic interactions between traits may slow selection responses owing to tradeoffs or when there is a low genetic variance for traits in the direction of selection³⁴. Strong

constraining interactions or a low level of genetic variance can therefore provide an indication of whether rapid evolutionary adaptation is likely, but this is difficult to measure in practice.

Incorporating demographics and gene flow

Some models that predict evolutionary responses to climate change include details about population size and shifting patterns of stabilizing selection on populations (Table 2), emphasizing the importance of demographic factors in limiting adaptive responses in populations^{47,48}. Large effective population sizes are required for maintaining genetic variation and evolutionary potential—typically a thousand rather than a hundred breeding individuals are required^{47,49}. These models have recently been extended to include plasticity; this relaxes the conditions under which extinction is inevitable unless the costs of plasticity are high⁵⁰. At small population sizes, demographic and environmental stochasticity will have a much larger impact on extinction probabilities than genetic variation⁴⁹, and demographic factors may lead to extinction even when populations have the requisite genetic variation to evolve⁵¹.

Demographic factors have been emphasized in models of asexual populations where mutational input enables adaptive responses and beneficial mutations (already present in populations or newly arisen) increase in frequency to rescue populations by providing new genetic variation (Table 2). To avoid extinction, mutation rate and population size need to be sufficient to maintain beneficial mutations and counter demographic effects^{10,52}. Populations need to consist of several hundred individuals for rescue through the introduction of new genetic variants to be effective when there is a sudden decrease in population size⁵³.

Gene flow among populations also influences the potential for evolutionary responses (Table 2), particularly at species margins that often represent sinks for dispersers from within the centre of species ranges. These dispersers may be maladapted to the conditions experienced at range margins, resulting in a fitness cost that prevents adaptation^{54,55}. This might constrain evolutionary responses, as recently suggested for barnacles⁵⁶. Gene flow may act in combination with other factors like trait interactions⁵⁷ and demographic effects⁵⁵ to constrain evolution.

Both demographic factors and gene flow can be incorporated into individual-based models of selection (Table 2). When these models are applied to elevation gradients in birch and Scots pine, they indicate that evolution is expected to keep up with a rapid rate of environmental change much more effectively when there is high mortality in established birch and pine trees, regardless of gene flow differences between the tree species⁵⁸. Given sufficient knowledge about patterns of selection, dispersal and population size, individual-based models can provide detailed predictions about the ability of species to track environmental change through evolution.

Lastly, statistical models help predict the likelihood of evolution on the basis of the rates at which phenotypic changes in traits have evolved in the past (Table 2). Phenotypic changes can be dissected into genetic and non-genetic components according to formulations of the Price equation²¹; the strength of any evolutionary shifts can then be used to predict evolutionary adaptation into the future.

Adaptation and the rate of environmental change

The single population models in Table 2 predict that extinction rather than evolution is likely if environmental conditions change too quickly. Evolutionary responses might only be sufficient to allow changes of a few per cent per generation⁴⁷ unless there are also plastic effects acting alongside evolution⁵⁰, whereas bird breeding dates have shifted almost 20% in the last few years¹⁹. This raises the question of whether long-lived organisms in particular can evolve to keep up with rapid climate change.

As a lag develops between environmental conditions and the optimum values of a trait⁴⁷, the intensity of directional selection on the trait will increase. This could theoretically lead to a faster evolutionary response. Using the breeder's equation (Box 1), stronger directional selection should increase the rate of response across a generation because the selection differential will increase if heritability remains constant.

Table 2 | Models applied to understand evolutionary responses (or lack of these responses) and distribution shifts under climate change

Trait, population	Model	Nature of constraints	Limitations
One trait, one population	Breeder's equation (Box 1), where the directional response to selection across a generation depends on the heritability of a trait and intensity of selection.	Arise when heritability (h^2) is low in the environment where selection occurs ³⁷ and/or selection is too strong ² .	Only applicable to one generation, although often applied to several generations. No interactions among traits, no stochasticity in demography or environment, indicates extent to which mean of trait can be changed under constant selection, h^2 .
Multiple traits, one population	Multivariate equivalent of breeder's equation (Box 1), where continuous directional selection affects the vector of linear selection gradients and also the genetic variance–covariance matrix if this changes with the environment.	Constraints to selection responses can be detected through a combination of approaches ³⁴ and occur when there is no genetic variance in the direction of selection, even in the absence of strong negative genetic correlations among traits.	Stochasticity in demography or environment not considered, limits only identified if relevant traits are included in variance–covariance matrix, which may alter across generations.
Multiple traits, one population	Various equations for describing how selection results in changes in traits that are connected to fitness (Box 1).	Link multiple traits to fitness, and then assess changes in traits when selection and interactions with other traits are taken into account ⁴³ .	As for multiple traits, one population.
One trait, one population (varying in size)	Models involving stabilizing selection with a shifting peak where fitness is normally distributed about a changing optimum value shifting at different rates (Box 1) ⁴⁷ , may include stochasticity ⁵¹ .	Indicates the rate of environmental change that can be countered through evolution, also highlights effects of a decrease in population size in limiting evolutionary change.	Depends on traits being under stabilizing selection, as for other single trait models.
One trait (showing plasticity), one population	Stabilizing selection with a shifting peak where fitness is normally distributed about a changing optimum value and plasticity occurs (Box 1), which may have a cost ⁵⁰ .	Alternative formulation for critical environmental rate that can be countered through evolution, highlights potential effects of plasticity on countering extinction unless there are large costs associated with these responses.	As for one trait, one population (varying size).
Beneficial mutation influences fitness, one population	Model of sudden decrease in growth rate of population following stress, countered by evolutionary rescue through beneficial mutation (Box 1) ¹⁰ .	Rescue depends on having a sufficiently high genomic mutation rate, population growth rate and population size.	Model is applicable to asexual microbial populations where model parameters can be estimated.
Marginal population and source population, unidirectional gene flow	Various models and simulations ^{54,55} of directional selection in marginal populations that predict potential for adaptation at margins under different levels of gene flow.	Intermediate levels of gene flow favour adaptation, ensuring presence of genetic variation but ensuring that dispersal load is not too high. Small population size and increasing fluctuations at margins make adaptation less likely ⁵⁵ .	Only applicable to a population spreading from periphery but constrained by migration load.
One trait (limiting fitness), multiple populations along gradient	Individual-based model on the basis of a mechanistic model of growth along an environmental gradient, with directional selection on trait depending on position along gradient ⁵⁸ .	Successful adaptation depends on patterns of dispersal and selection, whereas demographic factors influence efficiency of selection.	Only allows relative importance of dispersal and demographic factors to be considered when there is enough information on trait responses along gradients.
One trait, one population in the past	Price equation describing whether changes in trait values under climate change are due to selection or other factors including plastic responses to the environment ²¹ . Changes in average value of a trait over generations, decomposed into components due to selection and other factors, in particular environmental values.	The change in mean phenotypic trait value across a generation is given by $\Delta\bar{z} = \text{cov}(w/\bar{w}, z) + E[(w/\bar{w})\delta z]$, where the first term is the covariance between trait value (z) and relative fitness (w/\bar{w}), and reflects the effects of trait selection, whereas the second term reflects the change in a trait arising from a lack of fidelity of transmission across generations, with δz being the discrepancy between an individual and offspring trait. Constraints arise when there is a large discrepancy and the effects of selection are not passed on across generations.	Statistical model for describing changes rather than predicting evolutionary responses where parameter values are known.
One trait (limiting distributions), multiple populations across landscape	Mechanistic model that is spatially explicit with threshold evolutionary model based on breeder's equation (Box 1), and climate variables that set selection intensity varying geographically ⁷ .	Indicates potential of evolution to influence shifts in distribution under climate change, tests whether evolution has an impact on distribution and whether abundance shifts under climate change.	As for one trait, one population (limiting fitness), does not consider changes in population size or gene flow.

However, genetic variability available to selection might decrease rapidly under directional selection, particularly in small populations.

Under intense selection, trait changes are more likely to be due to genes with large effects, but major genes are often associated with deleterious effects⁵⁹, as illustrated by populations of the alga *Chlamydomonas reinhardtii* exposed to stressful conditions for 200 generations⁶⁰. These populations had lower growth rates when more intense selection favoured genes with larger effects on traits, whereas slower rates of evolution reduced the cost of adaptation. It is not clear how often intense selection through rapid climate change will be constrained by the pleiotropic effects of major genes (when the same gene(s) influence two or more traits). In nature, selection will fluctuate from generation to generation⁶¹, and selection responses will depend not only on the intensity of selection but also on the way environmental fluctuations are correlated⁶².

Experimental evolution and phylogenetic constraints

Experimental evolution (Table 1) is a popular approach for monitoring adaptive potential in populations, particularly in microbes. In these experiments species or communities are exposed to simulated conditions likely to be encountered in the future. Examples of this approach include populations of *Chlamydomonas* exposed to increasing concentrations of ambient CO₂ for several hundred generations⁶³ and experimental populations of the bacterium *Escherichia coli* exposed for 20,000 generations to intermediate or high temperatures⁶⁴. The *Chlamydomonas* experiments showed that increased CO₂ levels did not necessarily increase algal growth and carbon uptake, which are important for mitigation. The *E. coli* experiments showed that populations at 37 °C had improved performance in the range 27–39 °C, but decreased performance at lower and higher temperatures, indicating that at least some lines are likely to show

tradeoffs across different thermal conditions⁶⁵. These types of experiments can be used to test specific predictions about evolutionary constraints, which might arise owing to pleiotropy or DNA decay (accumulation of mutations in genes leading to loss of gene function). Pleiotropy probably contributed to the decline in performance of *E. coli* at thermal extremes, whereas genes underlying mechanisms for concentrating carbon may have decayed in *Chlamydomonas* under high CO₂.

If genetic limits in species and populations arise because of DNA decay or pleiotropy, both processes could reduce genetic variation, which in turn might provide a surrogate measure for evolutionary potential. *Drosophila* climate specialists from wet tropical rainforests lack genetic variance for desiccation and cold resistance presumably because alleles required have been lost through decay or pleiotropic selection, even though the species have genetic variation for other traits and neutral markers⁴². An understanding of the mechanisms that underlie physiological constraints—such as oxygen limitation in marine organisms—can point to reasons for evolutionary limits⁶⁶, and the presence of decay in the genes underlying these mechanisms could reflect a low evolutionary potential. For example, some marine organisms adapted to stable, cold and well-oxygenated environments lack functional genes coding for proteins and regulatory systems for dealing with warmer conditions⁶⁷. A fraction of climatically restricted species might never be able to adapt owing to such genetic limits. As information emerges on genes controlling variation in adaptive traits, genetic limits might eventually be identifiable from comparisons of species genomes³³ (Table 1).

Where phylogenetic information is available, evolutionary constraints can be studied across lineages. Related plant species including anemones, buttercups and mints that have been more prone to extinction in Thoreau's woods, Massachusetts, seem to lack variability in flowering time⁶⁸; this might reflect a low genetic variance for flowering time or its plasticity. Conversely, the flowering time of invasive species has been responsive to temperature, potentially contributing to their expansion under recent climate change⁶⁹.

Hybridization

As some populations and species spread under favourable climatic conditions, new contact zones arise between related lineages, leading to interspecific competition but also an increased likelihood of hybridization between taxa⁷⁰. Hybridization is often regarded as a negative outcome for conservation, both because diversity is lost when a species' genome is replaced, and because fitness declines following admixture⁷¹. Genome replacement is particularly likely in populations that are becoming small and peripheral under climate change, as in the case of peripheral wintergreen (*Pyrola minor*) populations facing extinction due to hybridization with more abundant *P. grandiflora*⁷².

However, hybridization can also facilitate evolutionary adaptation. Molecular evidence indicates that the expansion of species' climatic ranges can be a consequence of past hybridization⁷³. Moreover, hybridization can increase the evolutionary potential of populations by introducing genetic variation, as in Darwin's finches where interspecies hybridization has provided most of the genetic variance in morphology for adapting to changing conditions⁷⁴. Lastly, hybridization may facilitate adaptation to new environments when hybridizing species are initially adapted to different conditions⁷⁵. As species distributions shift under climate change and the incidence of hybridization increases, there may be some unexpected evolutionary consequences and even benefits when new variation is introduced into populations lacking adaptive potential.

Evolution in distribution modelling

With an increasing number of studies demonstrating the potential for evolutionary changes in populations (Table 1) and documenting its effects in natural populations through comparisons of traits or genetic markers¹¹, it seems appropriate to incorporate evolution into models that predict shifts in the distribution and abundance of species. This applies particularly to species with short generation times and large population sizes (that is, with the potential to evolve rapidly and maintain genetic

variation). At present, correlative distribution models, which have become particularly popular for predicting distribution shifts⁷⁶, consider correlations between species distributions and current climatic variables. These only include evolutionary potential to the extent that adaptive geographical variation affects the climate niche of species. However, the geographical ranges of introduced species commonly cover climate space outside areas in the native range, pointing to the importance of factors such as rapid evolution and species interactions in driving range expansion⁷⁷.

Evolutionary changes in traits should be incorporated into mechanistic models of species distributions. These models of energy and mass transfer for organisms and their microclimates start from first principles to identify the traits that limit distributions⁷⁸, which can then be allowed to evolve (Table 2). An example of this approach (Fig. 1) predicts changes in the distribution of the mosquito *Aedes aegypti* as a consequence of evolution in egg desiccation resistance⁷. In this model, evolution allows the species to become established in the city of Darwin under climate change, where it could then potentially transmit dengue fever. Without evolution, the species would fail to establish in Darwin. Mechanistic models have also been used to predict adaptive dynamics in plant populations⁵⁸.

Evolutionary components could be routinely built into mechanistic models to explore the potential impact of evolution on distribution shifts. Estimates of heritable variation, selection intensity and even demographic effects on genetic variance could be incorporated into these models.

Managing evolution

Because evolutionary processes are potentially important in dealing with climate change, evolutionary criteria should be considered in biodiversity management strategies^{3,50}. DNA decay in genes that are functionally important, low levels of genetic variation, or phyletic conservatism may point to groups of species being particularly susceptible to climate change because of a limited capacity for adaptation. These species groups should receive particular attention in conservation programs.

Many populations of threatened species are confined to habitat fragments with limited options for dispersal under climate change. The adaptive potential of these populations is probably already low if they are small and have inherently low levels of genetic variance for traits limiting their distribution. For these populations, managers should consider any process that might increase their evolutionary potential,

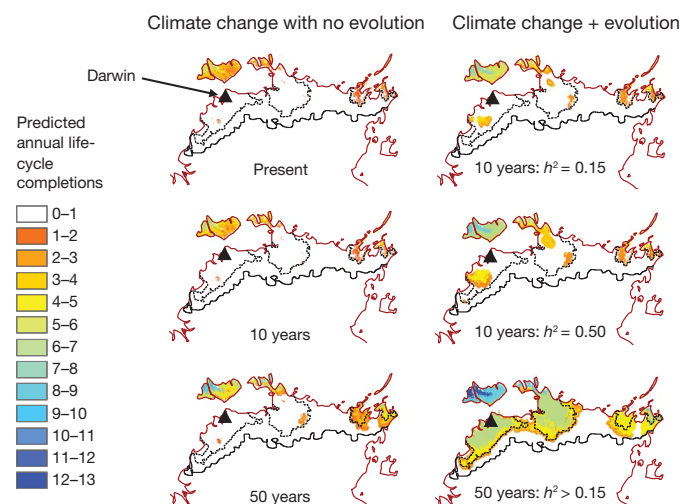


Figure 1 | Potential effect of evolution in egg desiccation resistance on life-cycle completions of the mosquito *Aedes aegypti* in the region around Darwin, Australia. Predictions are based on a mechanistic model of mosquito development with egg desiccation resistance evolving as a threshold trait determined by the length of time that eggs are exposed to air after water in breeding sites has evaporated. Modified with permission from ref. 7.

including improving connectedness with other populations even in the face of evidence for genetic distinctiveness⁷⁹.

Evolutionary potential could be used to identify populations and species that might be threatened in the future. Criteria based on patterns of local adaptation to climate gradients, probable levels of genetic variance now and into the future, and exposure to ongoing climate selection could all be used to identify species at risk. Models that incorporate the effects of climate change with evolutionary adaptation, demographics and stochastic variation could be developed to predict the likelihood of adaptation or extinction of populations⁵⁸. Gene flow among populations in fragments may need to be maintained at intermediate levels to maximize local adaptation⁵⁵.

By considering evolution, the likelihood that key plant and animal species will persist within landscapes under climate change can be increased. For revegetation projects, seed of key species could be sourced from climatically diverse areas to increase the future speed of adaptation^{80,81}, although there is a risk that this might increase disease transmission and in some cases produce outbreeding depression (where the fitness of genotypes from population crosses is low). Where these problems are minor, seed and pollen could be deliberately exchanged between populations as a way of ensuring that some genotypes adapted to future climatic conditions exist in populations⁸⁰. The genetic breadth of stock used in nurseries for propagating plants, fish and other groups could be expanded to capture evolutionary diversity in populations, including sourcing material from different climates.

When designating areas for reserves, areas containing high genetic variation across multiple species should be prioritized⁸². Areas for conservation should include climatic gradients where selection varies over small geographical areas, ensuring that genotypes adapted to different conditions are maintained⁸³. Environmental heterogeneity also provides a way of ensuring that phenotypic variability is maintained in species⁸⁴, increasing the adaptedness of populations as conditions change. Reserves could be designed or modified to maximize evolutionary processes, by connecting areas to enable gene flow, and allowing for ongoing selection in response to fire, drought, thermal stresses and changing species interactions.

Future directions

The literature is now replete with examples illustrating the power of evolution to generate rapid phenotypic changes and influence community dynamics. However, empirical and theoretical studies indicate that some populations will have difficulty in adapting to climate change because of low trait heritability, overriding environmental effects on phenotypes, and strong but fluctuating selection pressures. Far more empirical data are needed to test evolutionary potential across groups of species, including those sensitive to thermal extremes and intermittently dry conditions. Evolutionary models of climate change responses need to be extended to incorporate stochastic climatic conditions. Individual-based models can provide precise predictions about the level of environmental change to which populations might adapt and the evolutionary processes that facilitate adaptation, but these models require detailed information only available in a few cases. In the absence of such information for the majority of species and in light of the complex selection pressures likely to arise from climate change, it seems prudent to incorporate evolutionary considerations in predictive modelling, and to take a precautionary approach when conserving sensitive species with a low adaptive potential.

1. Parmesan, C. & Yohe, G. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**, 37–42 (2003).
2. Sinervo, B. *et al.* Erosion of lizard diversity by climate change and altered thermal niches. *Science* **328**, 894–899 (2010).
3. Williams, S. E., Shoo, L. P., Isaac, J. L., Hoffmann, A. A. & Langham, G. Towards an integrated framework for assessing the vulnerability of species to climate change. *PLoS Biol.* **6**, 2621–2626 (2008).
4. Hendry, A. P., Farrugia, T. J. & Kinnison, M. T. Human influences on rates of phenotypic change in wild animal populations. *Mol. Ecol.* **17**, 20–29 (2008).
5. Whitney, K. D. & Gabler, C. A. Rapid evolution in introduced species, 'invasive traits' and recipient communities: challenges for predicting invasive potential. *Divers. Distrib.* **14**, 569–580 (2008).

6. Carroll, S. P. Facing change: forms and foundations of contemporary adaptation to biotic invasions. *Mol. Ecol.* **17**, 361–372 (2008).
7. Kearney, M., Porter, W. P., Williams, C., Ritchie, S. & Hoffmann, A. A. Integrating biophysical models and evolutionary theory to predict climatic impacts on species' ranges: the dengue mosquito *Aedes aegypti* in Australia. *Funct. Ecol.* **23**, 528–538 (2009).
- Incorporated for the first time trait evolution into a mechanistic model, showing that predictions about species-distribution shifts under climate change can be inaccurate if evolution is ignored.**
8. Kanarek, A. & Webb, C. Allee effects, adaptive evolution, and invasion success. *Evol. Appl.* **3**, 122–135 (2010).
9. Klenner, W. & Arsenault, A. Ponderosa pine mortality during a severe bark beetle (Coleoptera: Curculionidae, Scolytinae) outbreak in southern British Columbia and implications for wildlife habitat management. *For. Ecol. Manage.* **258**, S5–S14 (2009).
10. Bell, G. & Collins, S. Adaptation, extinction and global change. *Evol. Appl.* **1**, 3–16 (2008).
- Using experimental evolution, demonstrates that evolutionary adaptation to climate change can influence the ability of species to take advantage of potentially favourable conditions arising from climate change, including the effects of CO₂ enrichment on growth rate.**
11. Bradshaw, W. E. & Holzapfel, C. M. Genetic response to rapid climate change: it's seasonal timing that matters. *Mol. Ecol.* **17**, 157–166 (2008).
12. Stillman, J. H. Causes and consequences of thermal tolerance limits in rocky intertidal porcelain crabs, genus *Petrolisthes*. *Integr. Comp. Biol.* **42**, 790–796 (2002).
13. Huey, R. B. *et al.* Why tropical forest lizards are vulnerable to climate warming. *Proc. R. Soc. B* **276**, 1939–1948 (2009).
14. Crispo, E. *et al.* The evolution of phenotypic plasticity in response to anthropogenic disturbance. *Evol. Ecol. Res.* **12**, 47–66 (2010).
15. Van Der Jeugd, H. P. *et al.* Keeping up with early springs: rapid range expansion in an avian herbivore incurs a mismatch between reproductive timing and food supply. *Glob. Change Biol.* **15**, 1057–1071 (2009).
16. Gozlan, R. E., Britton, J. R., Cowx, I. & Copp, G. H. Current knowledge on non-native freshwater fish introductions. *J. Fish Biol.* **76**, 751–786 (2010).
17. Forister, M. L. *et al.* Compounded effects of climate change and habitat alteration shift patterns of butterfly diversity. *Proc. Natl Acad. Sci. USA* **107**, 2088–2092 (2010).
18. Hadfield, J. D., Wilson, A. J., Garant, D., Sheldon, B. C. & Kruuk, L. E. B. The misuse of BLUP in ecology and evolution. *Am. Nat.* **175**, 116–125 (2010).
- Highlights the potential pitfalls associated with using quantitative genetic approaches to identify genetic changes in populations over time.**
19. Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A. & Merila, J. Climate change and evolution: disentangling environmental and genetic responses. *Mol. Ecol.* **17**, 167–178 (2008).
20. Ozgul, A. *et al.* Coupled dynamics of body mass and population growth in response to environmental change. *Nature* **466**, 482–485 (2010).
21. Ozgul, A. *et al.* The dynamics of phenotypic change and the shrinking sheep of St. Kilda. *Science* **325**, 464–467 (2009).
22. Etterson, J. R. Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. I. Clinal patterns of selection along an environmental gradient in the great plains. *Evolution* **58**, 1446–1458 (2004).
23. Linhart, Y. B. & Grant, M. C. Evolutionary significance of local genetic differentiation in plants. *Annu. Rev. Ecol. Syst.* **27**, 237–277 (1996).
24. Ayirinhac, A. *et al.* Cold adaptation in geographical populations of *Drosophila melanogaster*: phenotypic plasticity is more important than genetic variability. *Funct. Ecol.* **18**, 700–706 (2004).
25. Hoffmann, A. A., Shirriffs, J. & Scott, M. Relative importance of plastic vs genetic factors in adaptive differentiation: geographical variation for stress resistance in *Drosophila melanogaster* from eastern Australia. *Funct. Ecol.* **19**, 222–227 (2005).
26. Kingsolver, J. G., Massie, K. R., Ragland, G. J. & Smith, M. H. Rapid population divergence in thermal reaction norms for an invading species: breaking the temperature-size rule. *J. Evol. Biol.* **20**, 892–900 (2007).
27. Franks, S. J., Sim, S. & Weis, A. E. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proc. Natl Acad. Sci. USA* **104**, 1278–1282 (2007).
28. Pulido, F. & Berthold, P. Current selection for lower migratory activity will drive the evolution of residency in a migratory bird population. *Proc. Natl Acad. Sci. USA* **107**, 7341–7346 (2010).
29. Gardner, J. L., Heinsohn, R. & Joseph, L. Shifting latitudinal clines in avian body size correlate with global warming in Australian passerines. *Proc. R. Soc. B* **276**, 3845–3852 (2009).
30. Bradshaw, W. E., Armbruster, P. A. & Holzapfel, C. M. Fitness consequences of hibernation diapause in the pitcher-plant mosquito, *Wyeomyia smithii*. *Ecology* **79**, 1458–1462 (1998).
31. Ragland, G. J. & Kingsolver, J. G. Evolution of thermotolerance in seasonal environments: the effects of annual temperature variation and life-history timing in *Wyeomyia smithii*. *Evolution* **62**, 1345–1357 (2008).
32. Schmidt, P. S., Matzkin, L., Ippolito, M. & Eanes, W. F. Geographic variation in diapause incidence, life-history traits, and climatic adaptation in *Drosophila melanogaster*. *Evolution* **59**, 1721–1732 (2005).
33. Hoffmann, A. A. A genetic perspective on insect climate specialists. *Aust. J. Entomol.* **49**, 93–103 (2010).
34. Walsh, B. & Blows, M. W. Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. *Annu. Rev. Ecol. Syst.* **40**, 41–59 (2009).

35. van Asch, M., Tienderen, P. H., Holleman, L. J. M. & Visser, M. E. Predicting adaptation of phenology in response to climate change, an insect herbivore example. *Glob. Change Biol.* **13**, 1596–1604 (2007).
36. Maloney, S. K., Fuller, A. & Mitchell, D. Climate change: is the dark Soay sheep endangered? *Biol. Lett.* **5**, 826–829 (2009).
37. Hoffmann, A. A., Sorensen, J. G. & Loeschcke, V. Adaptation of *Drosophila* to temperature extremes: bringing together quantitative and molecular approaches. *J. Therm. Biol.* **28**, 175–216 (2003).
38. Jump, A. S. et al. Simulated climate change provokes rapid genetic change in the Mediterranean shrub *Fumana thymifolia*. *Glob. Change Biol.* **14**, 637–643 (2008).
39. Mimura, M. & Aitken, S. N. Local adaptation at the range peripheries of Sitka spruce. *J. Evol. Biol.* **23**, 249–258 (2010).
40. Loehr, J., Carey, J., O'Hara, R. B. & Hik, D. S. The role of phenotypic plasticity in responses of hunted thinhorn sheep ram horn growth to changing climate conditions. *J. Evol. Biol.* **23**, 783–790 (2010).
41. Teplitsky, C., Mills, J. A., Yarrall, J. W. & Merila, J. Indirect genetic effects in a sex-limited trait: the case of breeding time in red-billed gulls. *J. Evol. Biol.* **23**, 935–944 (2010).
42. Kellermann, V., van Heerwaarden, B., Sgrò, C. M. & Hoffmann, A. A. Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. *Science* **325**, 1244–1246 (2009).
43. Etterson, J. R. & Shaw, R. G. Constraint to adaptive evolution in response to global warming. *Science* **294**, 151–154 (2001).
44. Sheldon, B. C., Kruuk, L. E. B. & Merila, J. Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution* **57**, 406–420 (2003).
45. Tonsor, S. J. et al. Heat shock protein 101 effects in *A. thaliana*: genetic variation, fitness and pleiotropy in controlled temperature conditions. *Mol. Ecol.* **17**, 1614–1626 (2008).
46. Kristensen, T. N., Loeschcke, V. & Hoffmann, A. A. Can artificially selected phenotypes influence a component of field fitness? Thermal selection and fly performance under thermal extremes. *Proc. R. Soc. B* **274**, 771–778 (2007).
47. Lynch, M. & Lande, R. in *Biotic Interactions and Global Change* (eds Kareiva, P. M., Kingsolver, J. G. & Huey, R. B.) (Sinauer, 1993).
48. Burger, R. & Lynch, M. Evolution and extinction in a changing environment—a quantitative-genetic analysis. *Evolution* **49**, 151–163 (1995).
49. Willi, Y. & Hoffmann, A. A. Demographic factors and genetic variation influence population persistence under environmental change. *J. Evol. Biol.* **22**, 124–133 (2009).
50. Chevin, L.-M., Lande, R. & Mace, G. M. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biol.* **8**, e1000357 (2010).
- Review and advance in theory considering the effects of plasticity and evolutionary adaptation in combination.**
51. Gomulkiewicz, R. & Holt, R. D. When does evolution by natural selection prevent extinction? *Evolution* **49**, 201–207 (1995).
52. Orr, H. A. & Unckless, R. L. Population extinction and the genetics of adaptation. *Am. Nat.* **172**, 160–169 (2008).
53. Bell, G. & Gonzalez, A. Evolutionary rescue can prevent extinction following environmental change. *Ecol. Lett.* **12**, 942–948 (2009).
54. Kirkpatrick, M. & Barton, N. H. Evolution of a species' range. *Am. Nat.* **150**, 1–23 (1997).
55. Bridle, J. R., Polechova, J., Kawata, M. & Butlin, R. K. Why is adaptation prevented at ecological margins? New insights from individual-based simulations. *Ecol. Lett.* **13**, 485–494 (2010).
56. Dawson, M. N., Grosberg, R. K., Stuart, Y. E. & Sanford, E. Population genetic analysis of a recent range expansion: mechanisms regulating the poleward range limit in the volcano barnacle *Tetracita rubescens*. *Mol. Ecol.* **19**, 1585–1605 (2010).
57. Angert, A. L., Bradshaw, H. D. & Schemske, D. W. Using experimental evolution to investigate geographic range limits in Monkeyflowers. *Evolution* **62**, 2660–2675 (2008).
58. Kuparinen, A., Savolainen, O. & Schurr, F. M. Increased mortality can promote evolutionary adaptation of forest trees to climate change. *For. Ecol. Manage.* **259**, 1003–1008 (2010).
- Uses empirical data to parameterize an individual-based model of adaptation in two trees along a gradient and set the limits of adaptive genetic change.**
59. Orr, H. A. The population genetics of beneficial mutations. *Philos. Trans. R. Soc. B* **365**, 1195–1201 (2010).
60. Collins, S. & de Meaux, J. Adaptation to different rates of environmental change in *Chlamydomonas*. *Evolution* **63**, 2952–2965 (2009).
61. Grant, P. R. & Grant, B. R. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**, 707–711 (2002).
62. Björklund, M. et al. Quantitative trait evolution and environmental change. *PLoS ONE* **4**, e4521 (2009).
63. Collins, S. & Bell, G. Evolution of natural algal populations at elevated CO₂. *Ecol. Lett.* **9**, 129–135 (2006).
64. Cooper, V. S., Bennett, A. F. & Lenski, R. E. Evolution of thermal dependence of growth rate of *Escherichia coli* populations during 20,000 generations in a constant environment. *Evolution* **55**, 889–896 (2001).
65. Bennett, A. F. & Lenski, R. E. An experimental test of evolutionary trade-offs during temperature adaptation. *Proc. Natl Acad. Sci. USA* **104**, 8649–8654 (2007).
66. Pörtner, H. O. Oxygen- and capacity-limitation of thermal tolerance: a matrix for integrating climate-related stressor effects in marine ecosystems. *J. Exp. Biol.* **213**, 881–893 (2010).
67. Somero, G. N. The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers'. *J. Exp. Biol.* **213**, 912–920 (2010).
68. Willis, C. G., Ruhfel, B., Primack, R. B., Miller-Rushing, A. J. & Davis, C. C. Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. *Proc. Natl Acad. Sci. USA* **105**, 17029–17033 (2008).
69. Willis, C. G. et al. Favorable climate change response explains non-native species' success in Thoreau's Woods. *PLoS ONE* **5**, e8878 (2010).
70. Garroway, C. J. et al. Climate change induced hybridization in flying squirrels. *Glob. Change Biol.* **16**, 113–121 (2010).
71. Muhlfeld, C. C. et al. Hybridization rapidly reduces fitness of a native trout in the wild. *Biol. Lett.* **5**, 328–331 (2009).
72. Beatty, G. E., Philipp, M. & Provan, J. Unidirectional hybridization at a species' range boundary: implications for habitat tracking. *Divers. Distrib.* **16**, 1–9 (2010).
73. Besansky, N. J. et al. Semipermeable species boundaries between *Anopheles gambiae* and *Anopheles arabiensis*: evidence from multilocus DNA sequence variation. *Proc. Natl Acad. Sci. USA* **100**, 10818–10823 (2003).
74. Grant, P. R. & Grant, B. R. Conspecific versus heterospecific gene exchange between populations of Darwin's finches. *Philos. Trans. R. Soc. B* **365**, 1065–1076 (2010).
- Highlights the potential importance of hybridization in generating genetic variation for adapting to climate change.**
75. Donovan, L. A., Rosenthal, D. M., Sanchez-Velenosi, M., Rieseberg, L. H. & Ludwig, F. Are hybrid species more fit than ancestral parent species in the current hybrid species habitats? *J. Evol. Biol.* **23**, 805–816 (2010).
76. Elith, J. & Leathwick, J. R. Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Syst.* **40**, 677–697 (2009).
77. Gallagher, R. V., Beaumont, L. J., Hughes, L. & Leishman, M. R. Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *J. Ecol.* **98**, 790–799 (2010).
78. Kearney, M. & Porter, W. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecol. Lett.* **12**, 334–350 (2009).
79. Mace, G. M. & Purvis, A. Evolutionary biology and practical conservation: bridging a widening gap. *Mol. Ecol.* **17**, 9–19 (2008).
80. Broadhurst, L. M. et al. Seed supply for broadscale restoration: maximising evolutionary potential. *Evolutionary Applications* **1**, 587–597 (2008).
81. Jones, T. A. & Monaco, T. A. A role for assisted evolution in designing native plant materials for domesticated landscapes. *Front. Ecol. Environ.* **7**, 541–547 (2009).
82. Vandergast, A. G., Bohonak, A. J., Hathaway, S. A., Boys, J. & Fisher, R. N. Are hotspots of evolutionary potential adequately protected in southern California? *Biol. Conserv.* **141**, 1648–1664 (2008).
83. Moritz, C., Hoskin, C., Graham, C. H., Hugall, A. & Moussalli, A. in *Phylogeny and Conservation* (eds Purvis, A., Gittleman, J. L. & Brooks, T.) 243–267 (Cambridge Univ. Press, 2009).
84. Watters, J. V., Lema, S. C. & Nevitt, G. A. Phenotype management: a new approach to habitat restoration. *Biol. Conserv.* **112**, 435–445 (2003).
85. Dahlhoff, E. P. et al. Effects of temperature on physiology and reproductive success of a montane leaf beetle: implications for persistence of native populations enduring climate change. *Physiol. Biochem. Zool.* **81**, 718–732 (2008).
86. Visser, M. E. Keeping up with a warming world: assessing the rate of adaptation to climate change. *Proc. R. Soc. B* **275**, 649–659 (2008).
87. Van Doorslaer, W., Stoks, R., Duvivier, C., Bednarska, A. & De Meester, L. Population dynamics determine genetic adaptation to temperature in *Daphnia*. *Evolution* **63**, 1867–1878 (2009).
88. Collins, S. & Bell, G. Phenotypic consequences of 1,000 generations of selection at elevated CO₂ in a green alga. *Nature* **431**, 566–569 (2004).
89. Larsen, P. F. et al. Adaptive differences in gene expression in European flounder (*Platichthys flesus*). *Mol. Ecol.* **16**, 4674–4683 (2007).
90. Michalski, S. G. et al. Evidence for genetic differentiation and divergent selection in an autotetraploid forage grass (*Arrhenatherum elatius*). *Theor. Appl. Genet.* **120**, 1151–1162 (2010).

Acknowledgements A.A.H. was supported by an ARC Australian Laureate Fellowship. C.M.S. was supported by an ARC Australian Research Fellowship. A.A.H. and C.M.S. were also supported by funding provided by the Commonwealth Environment Research Facility. We thank P. Griffin, A. Miller and C. Robin for comments that improved the manuscript.

Author Contributions A.A.H. developed the framework for the review, both authors contributed sections.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.A.H. (ary@unimelb.edu.au).

An open-system quantum simulator with trapped ions

Julio T. Barreiro^{1*}, Markus Müller^{2,3*}, Philipp Schindler¹, Daniel Nigg¹, Thomas Monz¹, Michael Chwalla^{1,2}, Markus Hennrich¹, Christian F. Roos^{1,2}, Peter Zoller^{2,3} & Rainer Blatt^{1,2}

The control of quantum systems is of fundamental scientific interest and promises powerful applications and technologies. Impressive progress has been achieved in isolating quantum systems from the environment and coherently controlling their dynamics, as demonstrated by the creation and manipulation of entanglement in various physical systems. However, for open quantum systems, engineering the dynamics of many particles by a controlled coupling to an environment remains largely unexplored. Here we realize an experimental toolbox for simulating an open quantum system with up to five quantum bits (qubits). Using a quantum computing architecture with trapped ions, we combine multi-qubit gates with optical pumping to implement coherent operations and dissipative processes. We illustrate our ability to engineer the open-system dynamics through the dissipative preparation of entangled states, the simulation of coherent many-body spin interactions, and the quantum non-demolition measurement of multi-qubit observables. By adding controlled dissipation to coherent operations, this work offers novel prospects for open-system quantum simulation and computation.

Every quantum system is inevitably coupled to its surrounding environment. Significant progress has been made in isolating systems from their environment and coherently controlling the dynamics of several qubits^{1–4}. These achievements have enabled the realization of high-fidelity quantum gates and the implementation of small-scale quantum computing and communication devices, as well as the measurement-based probabilistic preparation of entangled states in atomic^{5,6}, photonic⁷, NMR⁸ and solid-state set-ups^{9–11}. In particular, successful demonstrations of quantum simulators^{12,13}, which allow one to mimic and study the dynamics of complex quantum systems, have been reported¹⁴.

In contrast, controlling the more general dynamics of open systems amounts to engineering both the Hamiltonian time evolution of the system as well as the coupling to the environment. In previous work^{15–18}, controlled decoherence has been used to systematically study the detrimental effects of decoherence on many-body or multi-qubit open systems. The ability to design dissipation can, however, be a useful resource, as in the context of the preparation of a desired entangled state from an arbitrary initial state^{19–21}, and in the closely related fields of dissipative quantum computation²² and quantum memories²³. It also enables the preparation and manipulation of many-body states and quantum phases²⁰, and provides an enhanced sensitivity in precision measurements²⁴. In particular, by combining suitably chosen coherent and dissipative operations, one can engineer the system–environment coupling, thus generalizing the concept of Hamiltonian quantum simulation to open quantum systems^{13,25}.

Here we provide an experimental demonstration of a toolbox of coherent and dissipative multi-qubit manipulations to control the dynamics of open systems. In a string of trapped ions, each ion encoding a qubit, we subdivide the qubits into ‘system’ and ‘environment’. The system–environment coupling is then engineered through the universal set of quantum operations available in ion-trap quantum computers^{26,27}, whereas the environment ion is coupled to

the dissipative bath of vacuum modes of the radiation field via optical pumping. Following ref. 22 (see also ref. 28), these quantum resources provide a complete toolbox to engineer general Markovian open-system dynamics in our multi-qubit system^{25,29}.

We first illustrate this engineering by dissipatively preparing a Bell state in a 2+1 ion system (that is, two system ions and one ancilla ion), such that an initially fully mixed state is pumped into a given Bell state. Similarly, with 4+1 ions, we also dissipatively prepare a four-qubit Greenberger–Horne–Zeiling (GHZ) state, which can be regarded as a minimal instance of Kitaev’s toric code³⁰. Besides the dissipative elements, we show coherent *n*-body interactions by implementing the fundamental building block for four-spin interactions. In addition, we demonstrate a readout of *n*-particle observables in a non-destructive way with a quantum-non-demolition (QND) measurement of a four-qubit stabilizer operator. We conclude by outlining future perspectives and implications of the present work for quantum information processing and simulation, as well as open-system quantum control scenarios including feedback²⁵.

Open-system dynamics and Bell-state pumping

The dynamics of an open quantum system *S* coupled to an environment *E* can be described by the unitary transformation $\rho_{SE} \mapsto U \rho_{SE} U^\dagger$, with ρ_{SE} the joint density matrix of the composite system *S* + *E*. Thus, the reduced density operator of the system will evolve as $\rho_S = \text{Tr}_E(U \rho_{SE} U^\dagger)$. The time evolution of the system can also be described by a completely positive Kraus map

$$\rho_S \mapsto \mathcal{E}(\rho_S) = \sum_k E_k \rho_S E_k^\dagger \quad (1)$$

with E_k operation elements satisfying $\sum_k E_k^\dagger E_k = 1$, and initially uncorrelated system and environment³¹. If the system is decoupled from the environment, the general map (1) reduces to $\rho_S \mapsto U_S \rho_S U_S^\dagger$, with U_S the unitary time evolution operator acting only on the system.

¹Institut für Experimentalphysik, Universität Innsbruck, Technikerstrasse 25, 6020 Innsbruck, Austria. ²Institut für Quantenoptik und Quanteninformation, Österreichische Akademie der Wissenschaften, Technikerstrasse 21A, 6020 Innsbruck, Austria. ³Institut für Theoretische Physik, Universität Innsbruck, Technikerstrasse 25, 6020 Innsbruck, Austria.

*These authors contributed equally to this work.

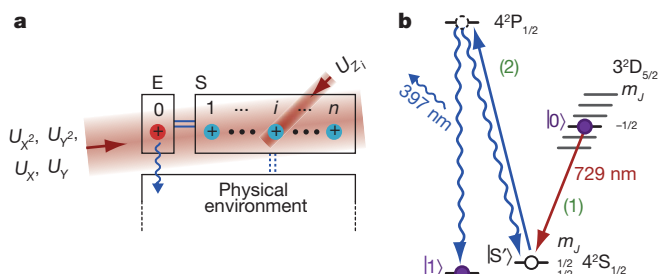


Figure 1 | Experimental tools for the simulation of open quantum systems with ions. **a**, The coherent component is realized by collective ($U_X, U_Y, U_{X^2}, U_{Y^2}$) and single-qubit operations (U_Z) on a string of $^{40}\text{Ca}^+$ ions which consists of the environment qubit (ion 0) and the system qubits (ions 1 to n). Coherent operations on S and E, combined with a controllable dissipative mechanism involving spontaneous emission of a photon from the environment ion, allow one to tailor the coupling of the system qubits to an artificial environment. This should be contrasted to the residual, detrimental coupling of the system (and environment) ions to their physical environment. **b**, The dissipative mechanism on the ancilla qubit is realized in the two steps shown on the Zeeman-split $^{40}\text{Ca}^+$ levels by (1) a coherent transfer of the population from $|0\rangle$ to $|S'\rangle$ (brown arrow) and (2) an optical pumping to $|1\rangle$ after a transfer to the $4^2P_{1/2}$ state by a circularly-polarized laser at 397 nm (represented by a blue straight arrow).

Control of both coherent and dissipative dynamics is then achieved by finding corresponding sequences of maps (1) specified by sets of operation elements $\{E_k\}$ and engineering these sequences in the laboratory. In particular, for the example of dissipative quantum-state preparation, pumping to an entangled state $|\psi\rangle$ reduces to implementing appropriate sequences of dissipative maps. These maps are chosen to drive the system to the desired target state irrespective of its initial state. The resulting dynamics have then the pure state $|\psi\rangle$ as the unique attractor, $\rho_S \mapsto |\psi\rangle\langle\psi|$. In quantum optics and atomic physics, the techniques of optical pumping and laser cooling are successfully used for the dissipative preparation of quantum states, although on a single-particle level. The engineering of dissipative maps for the preparation of entangled states can be seen as a generalization of this concept of pumping and cooling in driven dissipative systems to a many-particle context. To be concrete, we focus on dissipative preparation of stabilizer states, which represent a large family of entangled states, including graph states and error-correcting codes³².

We start by outlining the concept of Kraus map engineering for the simplest non-trivial example of ‘pumping’ a system of two qubits into a Bell state. The Hilbert space of two qubits is spanned by the four Bell states defined as $|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$ and $|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$. Here, $|0\rangle$ and $|1\rangle$ denote the computational basis of each qubit, and we use the short-hand notation $|00\rangle = |0\rangle_1|0\rangle_2$, for example. These maximally entangled states are stabilizer states: the Bell state $|\Phi^+\rangle$, for instance, is said to be stabilized by the two stabilizer operators Z_1Z_2 and X_1X_2 , where X and Z denote the usual Pauli matrices, as it is the only two-qubit state that is an eigenstate of eigenvalue +1 of these two commuting observables, that is, $Z_1Z_2|\Phi^+\rangle = |\Phi^+\rangle$ and $X_1X_2|\Phi^+\rangle = |\Phi^+\rangle$. In fact, each of the four Bell states is uniquely determined as an eigenstate with eigenvalues ± 1 with respect to Z_1Z_2 and X_1X_2 . The key idea of pumping is that we can achieve dissipative dynamics which pump the system into a particular Bell state, for example $\rho_S \mapsto |\Psi^-\rangle\langle\Psi^-|$, by constructing two dissipative maps, under which the two qubits are irreversibly transferred from the +1 into the -1 eigenspaces of Z_1Z_2 and X_1X_2 .

The dissipative maps are engineered with the aid of an ancilla ‘environment’ qubit^{25,33} and a quantum circuit of coherent and dissipative operations. The form and decomposition of these maps into basic operations are discussed in Box 1. The pumping dynamics are determined by the probability of pumping from the +1 into the -1 stabilizer eigenspaces, which can be directly controlled by varying the parameters in the employed gate operations. For pumping with unit probability ($p = 1$), the two qubits reach the target Bell state—regardless of their initial state—after only one pumping cycle, that is, by a single application

of each of the two maps. In contrast, when the pumping probability is small ($p \ll 1$), the process can be regarded as the infinitesimal limit of the general map (1). In this case, the system dynamics under a repeated application of the pumping cycle are described by a master equation³⁴:

$$\dot{\rho}_S = -i[H_S, \rho_S] + \sum_k \left(c_k \rho_S c_k^\dagger - \frac{1}{2} c_k^\dagger c_k \rho_S - \rho_S \frac{1}{2} c_k^\dagger c_k \right) \quad (2)$$

Here H_S is a system Hamiltonian, and c_k are Lindblad operators reflecting the system–environment coupling. For the purely dissipative maps discussed here, $H_S = 0$. Quantum jumps from the +1 into the -1 eigenspace of Z_1Z_2 and X_1X_2 are mediated by a set of two-qubit Lindblad operators (see Box 1 for details); here the system reaches the target Bell state asymptotically after many pumping cycles.

Experimental Bell-state pumping

The dissipative preparation of n -particle entangled states is realized in a system of $n + 1$ $^{40}\text{Ca}^+$ ions confined to a string by a linear Paul trap and cooled to the ground state of the axial centre-of-mass mode³⁵. For each ion, the internal electronic Zeeman levels $D_{5/2}(m = -1/2)$ and $S_{1/2}(m = -1/2)$ encode the logical states $|0\rangle$ and $|1\rangle$ of a qubit. For coherent operations, a laser at a wavelength of 729 nm excites the quadrupole transition connecting the qubit states ($S_{1/2} \leftrightarrow D_{5/2}$). A broad beam of this laser couples to all ions (Fig. 1a) and realizes the collective single-qubit gate $U_X(\theta) = \exp(-i\frac{\theta}{2}\sum_i X_i)$ as well as a Mølmer-Sørensen³⁶ (MS) entangling operation $U_{X^2}(\theta) = \exp(-i\frac{\theta}{4}(\sum_i X_i)^2)$ when using a bichromatic light field (θ is controlled by the intensity and length of the laser pulses). Shifting the optical phase of the drive field by $\pi/2$ exchanges X_i by Y_i in these operations. As a figure of merit of our entangling operation, we can prepare 3 (5) qubits in a GHZ state with 98% (95%) fidelity³⁷. These collective operations form a universal set of gates when used in conjunction with single-qubit rotations $U_{Z_i}(\theta) = \exp(-i\frac{\theta}{2}Z_i)$, which are realized by an off-resonant laser beam that can be adjusted to focus on any ion.

For engineering dissipation, the key element of the mapping steps, shown as (i) and (iii) in Box 1, is a single MS operation. The two-qubit gate, step (ii), is realized by a combination of collective and single-qubit operations. The dissipative mechanism, step (iv), is here carried out on the ancilla qubit by a reinitialization into $|1\rangle$, as shown in Fig. 1b. Another dissipative process (P.S. *et al.*, manuscript in preparation) can be used to prepare the system qubits in a completely mixed state by the transfer $|0\rangle \rightarrow (|0\rangle + |S'\rangle)/\sqrt{2}$ followed by optical pumping of $|S'\rangle$ into $|1\rangle$, where $|S'\rangle$ is the electronic level $S_{1/2}(m = 1/2)$.

Qubit read-out is accomplished by fluorescence detection on the $S_{1/2} \leftrightarrow P_{1/2}$ transition. The ancilla qubit can be measured without affecting the system qubits by applying hiding pulses that shelve the system qubits in the $D_{5/2}$ state manifold during fluorescence detection³⁸.

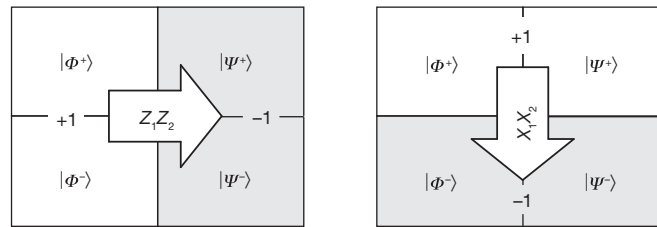
We use these tools to implement up to three Bell-state pumping cycles on a string of 2 + 1 ions. Starting with the two system qubits in a completely mixed state, we pump towards the Bell state $|\Psi^-\rangle$. Each pumping cycle is accomplished with a sequence of eight entangling operations, four collective unitaries and six single-qubit operations (see Supplementary Information). The pumping dynamics are probed by quantum state tomography of the system qubits after every half cycle. The reconstructed states are then used to map the evolution of the Bell-state populations.

In a first experiment, we set the pumping probability at $p = 1$ to observe deterministic pumping, and we obtain the Bell-state populations shown in Fig. 2a. As expected, the system reaches the target state after the first pumping cycle. Regardless of experimental imperfections, the target state population is preserved under the repeated application of further pumping cycles and reaches up to 91(1)% (all numbers in parentheses denote 1σ confidence intervals) after 1.5

BOX 1

Engineering dissipative open-system dynamics

Dissipative dynamics that pump two qubits from an arbitrary initial state into the Bell state $|\Psi^-\rangle$ are realized by two maps that generate pumping from the $+1$ into the -1 eigenspaces of the stabilizer operators Z_1Z_2 and X_1X_2 :



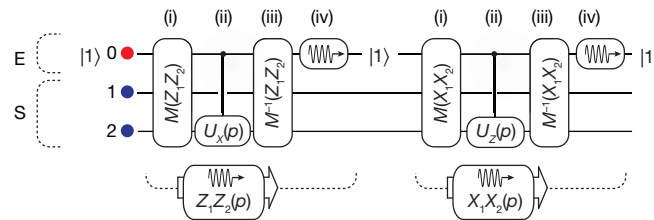
For Z_1Z_2 , the dissipative map pumping into the -1 eigenspace is $\rho_S \mapsto \mathcal{E}(\rho_S) = E_1 \rho_S E_1^\dagger + E_2 \rho_S E_2^\dagger$ with

$$E_1 = \sqrt{p} X_2 \frac{1}{2} (1 + Z_1 Z_2)$$

$$E_2 = \frac{1}{2} (1 - Z_1 Z_2) + \sqrt{1-p} \frac{1}{2} (1 + Z_1 Z_2)$$

The map's action as a unidirectional pumping process can be seen as follows. As the operation element E_1 contains the projector $\frac{1}{2}(1 + Z_1 Z_2)$ onto the $+1$ eigenspace of $Z_1 Z_2$, the spin flip X_2 can then convert $+1$ into -1 eigenstates of $Z_1 Z_2$; for example, $|\Phi^+\rangle \mapsto |\Psi^+\rangle$. In contrast, the -1 eigenspace of $Z_1 Z_2$ is left invariant. In the limit $p \ll 1$, the repeated application of this map reduces the process to a master equation with Lindblad operator $c = \frac{1}{2} X_2 (1 + Z_1 Z_2)$.

We implement the two dissipative maps by quantum circuits of three unitary operations (i)–(iii) and a dissipative step (iv). Both maps act on the two system qubits S and an ancilla which plays the role of the environment E .

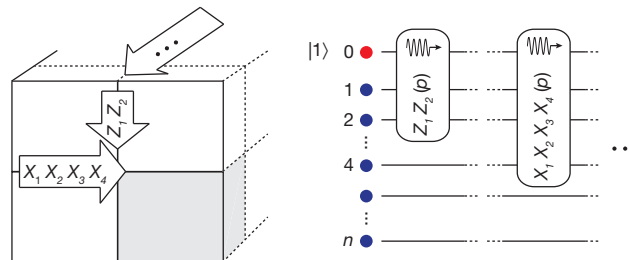


Pumping $Z_1 Z_2$ proceeds as follows: (i) Information about whether the system is in the $+1$ or -1 eigenspace of $Z_1 Z_2$ is mapped by $M(Z_1 Z_2)$ onto the logical states $|0\rangle$ and $|1\rangle$ of the ancilla (initially in $|1\rangle$). (ii) A controlled gate $C(p)$ converts $+1$ into -1 eigenstates by flipping the state of the second qubit with probability p , where

$$C(p) = |0\rangle\langle 0|_0 \otimes U_{X_2}(p) + |1\rangle\langle 1|_0 \otimes 1$$

with $U_{X_2}(p) = \exp(i\alpha X_2)$ and α controlling the pumping probability $p = \sin^2 \alpha$. (iii) The initial mapping is inverted by $M^{-1}(Z_1 Z_2)$. At this stage, in general, the ancilla and system qubits are entangled. (iv) The ancilla is dissipatively reset to $|1\rangle$, which carries away entropy to 'cool' the two system qubits. The second map for pumping into the -1 eigenspace of $X_1 X_2$ is obtained from interchanging the roles of X and Z above.

The engineering of dissipative maps can be readily generalized to systems of more qubits. As an example, dissipative preparation of n -qubit stabilizer states can be realized by a sequence of n dissipative maps (for example, for $Z_1 Z_2$ and $X_1 X_2 X_3 X_4$ pumping), which are implemented in analogy to the quantum circuits for Bell state pumping discussed above.



cycles (ideally 100%). In a second experiment, aimed at simulating master-equation dynamics, the probability is set at $p = 0.5$ to probe probabilistic pumping dynamics. The target state is then approached asymptotically (Fig. 2b). After pumping the system for three cycles with $p = 0.5$, up to 73(1)% of the initially mixed population is pumped into the target state (ideally 88%). To achieve Bell-state pumping in the limit of $p \ll 1$, the gate fidelities need to be raised closer to one because close to the stationary state of the dynamics, the pumping probability p per step for populating the target state competes directly

with loss processes at a rate ε . Such losses are associated with gate errors and lead to a steady state loss of fidelity that scales as $\propto \varepsilon/p$ (see Supplementary Information for further details).

In order to completely characterize the Bell-state pumping process, we also perform a quantum process tomography³¹. As an example, the reconstructed process matrix for $p = 1$ after 1.5 cycles (Fig. 2c) has a Jamiolkowski process fidelity³⁹ of 87.0(7)% with the ideal dissipative process $\rho_S \mapsto |\Psi^-\rangle\langle\Psi^-|$ which maps an arbitrary state of the system into the Bell state $|\Psi^-\rangle$.

Four-qubit stabilizer pumping

The engineering of the system–environment coupling, as demonstrated by Bell-state pumping above, can be readily extended to larger n -qubit open quantum systems. We illustrate such an engineering experimentally with the dissipative preparation of a four-qubit GHZ state $(|0000\rangle + |1111\rangle)/\sqrt{2}$. This state is uniquely characterized as the simultaneous eigenstate of the four stabilizers Z_1Z_2 , Z_2Z_3 , Z_3Z_4 and $X_1X_2X_3X_4$, all with eigenvalue $+1$ (Fig. 3a). Therefore, the pumping dynamics into the GHZ state are realized by four consecutive dissipative steps, each pumping the system into the $+1$ eigenspaces of the four stabilizers. In a system of $4+1$ ions, we implement such pumping dynamics in a manner analogous to the Bell-state pumping sequence. Here, however, the circuit decomposition of one pumping cycle involves 16 five-ion entangling operations, 20 collective unitaries and 34 single-qubit operations; further details are given in Supplementary Information.

In order to observe this deterministic pumping process into the GHZ state, we begin by preparing the system ions in a completely mixed state. The evolution of the state of the system after each pumping step is characterized by quantum state tomography. The reconstructed density matrices shown in Fig. 3b for the initial and subsequent states arising in each step have a fidelity (in %), or state overlap⁴⁰, with the expected states of $\{79(2), 89(1), 79.7(7), 70.0(7), 55.8(4)\}$ (the final state is genuinely multi-partite entangled⁴¹); see Supplementary Information for further details. The pumping dynamics is clearly reflected by the measured expectation values of the stabilizers Z_iZ_j ($ij = 12, 23, 34, 14$) and $X_1X_2X_3X_4$ at each step, as shown in Fig. 3c.

Although the simulation of a master equation strictly requires small pumping probabilities, we perform an exploratory study as follows. We implement up to five consecutive $X_1X_2X_3X_4$ -stabilizer pumping steps with two probabilities $p = 1$ and 0.5 , for the initial state $|1111\rangle$. The measured expectation values of all relevant stabilizers for pumping with $p = 1$ are shown in Fig. 3d. After the first step, the stabilizer $X_1X_2X_3X_4$ reaches an expectation value of $-0.68(1)$; after the second step and up to the fifth step, it is preserved at $-0.72(1)$ regardless of experimental imperfections.

For $X_1X_2X_3X_4$ -stabilizer pumping with $p = 0.5$, the four-qubit expectation value increases at each step and asymptotically approaches $-0.54(1)$ (ideally -1 ; fit shown in Fig. 3d). A state tomography after

each pumping step yields fidelities (in %) with the expected GHZ state of $\{53(1), 50(1), 49(1), 44(1), 41(1)\}$. From the reconstructed density matrices, we determine that the states generated after one to three cycles are genuinely multi-partite entangled⁴¹.

Coherent four-particle interactions

The coupling of the system to an ancilla particle, as used above for the engineering of dissipative dynamics, can also be harnessed to mediate effective coherent n -body interactions between the system qubits^{31,33}. The demonstration of a toolbox for open-system quantum simulation is thus complemented by adding unitary maps $\rho_S \mapsto U_S \rho_S U_S^\dagger$ to the dissipative elements described above. Here, $U_S = \exp(-i\tau H_S)$ is the unitary time evolution operator for a time step τ , which is generated by a system Hamiltonian H_S . In contrast to the recent achievements^{42,43} of small-scale analogue quantum simulators based on trapped ions, where two-body spin Hamiltonians have been engineered directly⁴⁴, here we pursue a gate-based implementation following the concept of Lloyd's digital quantum simulator¹³, where the time evolution is decomposed into a sequence of coherent (and dissipative) steps.

In particular, the available gate operations enable a simulation of n -body spin interactions, which we illustrate by implementing time dynamics of a four-body Hamiltonian $H_S = -gX_1X_2X_3X_4$. This example is motivated by the efforts to experimentally realize Kitaev's toric code Hamiltonian³⁰, which is a sum of commuting four-qubit stabilizer operators representing four-body spin interactions. This paradigmatic model belongs to a whole class of spin systems, which have been discussed in the context of topological quantum computing and quantum phases exhibiting topological order⁴⁵.

The elementary unitary operation U_S can be realized in three steps by the circuit shown in Fig. 4a. (i) As in the stabilizer pumping above, an operation $M(X_1X_2X_3X_4)$, here realized by an entangling MS gate $U_{X^2}(\pi/2)$, coherently maps the information about whether the four system spins are in the $+1(-1)$ eigenspace of $X_1X_2X_3X_4$ onto the internal states $|0\rangle$ and $|1\rangle$ of the ancilla qubit. (ii) Owing to this mapping, effectively all $+1(-1)$ eigenstates acquire a phase $\beta/2(-\beta/2)$ by a subsequent single-qubit rotation $U_Z(\beta)$ on the ancilla ion. The simulation time step τ is related to the phase by $\beta = 2g\tau$. (iii) After the initial mapping is inverted by a second MS gate $U_{X^2}(\pi/2)$, the ancilla qubit returns to its initial state and decouples from the four system qubits,

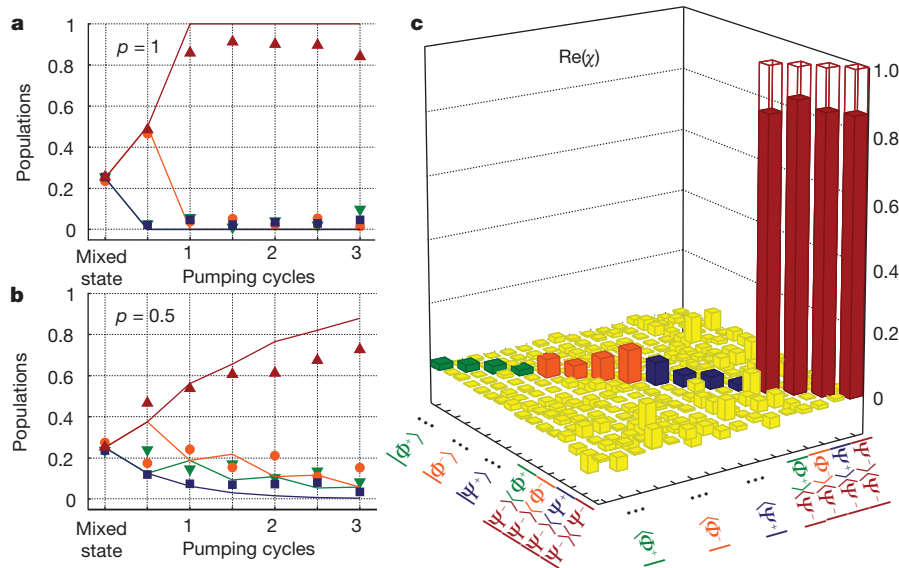


Figure 2 | Experimental signatures of Bell-state pumping. **a, b**, Evolution of the Bell-state populations $|\Phi^+\rangle$ (green down triangles), $|\Phi^-\rangle$ (orange circles), $|\Psi^+\rangle$ (blue squares) and $|\Psi^-\rangle$ (red up triangles) of an initially mixed state under a pumping process with probability $p = 1$ or deterministic (**a**) and $p = 0.5$ (**b**). Error bars, not shown, are smaller than 2% (1σ). **c**, Reconstructed process matrix χ (real part), displayed in the Bell-state basis, describing the

deterministic pumping of the two ions after one-and-a-half cycles. The ideal process mapping any input state into the state $|\Psi^-\rangle$ has as non-zero elements only the four transparent bars shown. The imaginary elements of χ , ideally all zero, have an average magnitude of 0.004 and a maximum of 0.03. The uncertainties in the elements of the process matrix are smaller than 0.01 (1σ). The colours of tickmarks and bars follow the colours used in **a** and **b**.

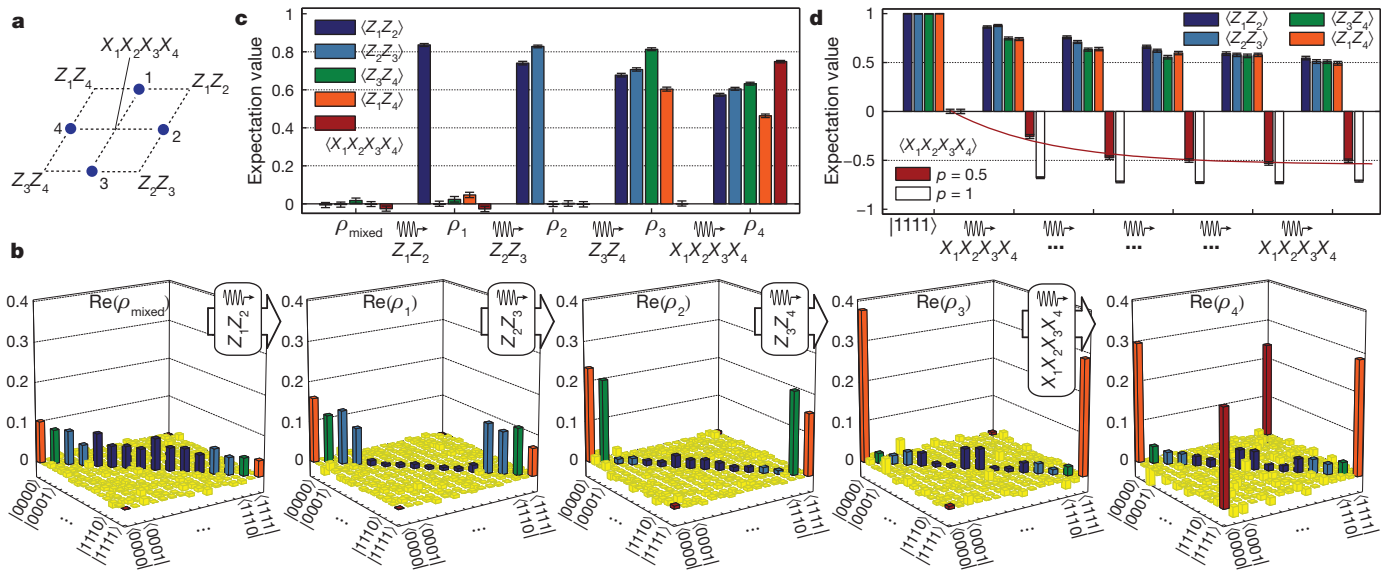


Figure 3 | Experimental signatures of four-qubit stabilizer pumping.

a, Diagram of the four system qubits to be pumped into the GHZ state $(|0000\rangle + |1111\rangle)/\sqrt{2}$, which is uniquely characterized as the simultaneous eigenstate with eigenvalue +1 of the shown stabilizers. **b**, Reconstructed density matrices (real part) of the initial mixed state ρ_{mixed} and subsequent states $\rho_{1,2,3,4}$ after sequentially pumping the stabilizers Z_1Z_2 , Z_2Z_3 , Z_3Z_4 and $X_1X_2X_3X_4$. Populations in the initial mixed state with qubits i and j antiparallel, or in the -1 eigenspace of the Z_iZ_j stabilizer, disappear after pumping this stabilizer into the $+1$ eigenspace. For example, populations in dark blue disappear after Z_1Z_2 -stabilizer pumping. A final pumping of the stabilizer $X_1X_2X_3X_4$ builds up the

which in turn have evolved according to U_S . This compact sequence makes the simulation of n -body interactions experimentally efficient. Here, the use of global MS gates conveniently bundles the effect of several operations (M.M., K. Hammerer, Y. Zhou, C.F.R. and P.Z., manuscript in preparation) which arise in alternative circuit decompositions based on two-qubit gates³¹.

In an experiment carried out with 4+1 ions, we apply U_S for different values of τ to the system ions initially prepared in $|1111\rangle$. We observe coherent oscillations in the subspace spanned by $|0000\rangle$ and $|1111\rangle$, as shown in Fig. 4b. We characterize our implementation of U_S by comparing the expected and measured states, determined by quantum state tomography, for each value of τ . The fidelity between the expected and measured states is on average 85(2)%.

QND measurement of a four-qubit stabilizer

Our toolbox for quantum simulation of open systems is extended by the possibility of reading out n -body observables in a non-destructive

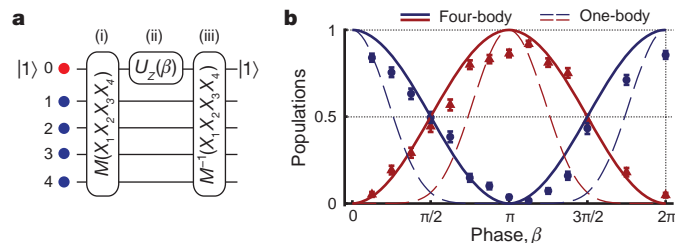


Figure 4 | Coherent simulation of four-body spin interactions. **a**, The elementary building block for the simulation of coherent evolution $U_S = \exp(-i\tau H_S)$ corresponding to the four-body Hamiltonian $H_S = -gX_1X_2X_3X_4$ ($\beta = 2g\tau$). **b**, Experimentally measured populations in state $|0000\rangle$ (up triangles) and $|1111\rangle$ (circles) as a function of β for a single application of U_S to the initial state $|1111\rangle$ of the four system qubits (error bars, $\pm 1\sigma$). The solid lines show the ideal behaviour. For comparison, the dashed lines indicate these populations for simultaneous single-qubit (one-body) oscillations, each driven by the rotation $\exp(-i\frac{\beta}{2}X_i)$.

coherence between $|0000\rangle$ and $|1111\rangle$, shown as red bars in the density matrix of ρ_4 . **c**, Measured expectation values of the relevant stabilizers; ideally, non-zero expectation values have a value of +1. **d**, Evolution of the measured expectation values of the relevant stabilizers for repeatedly pumping an initial state $|1111\rangle$ with probability $p = 0.5$ into the -1 eigenspace of the stabilizer $X_1X_2X_3X_4$. The incremental pumping is evident by the red line fitted to the pumped stabilizer expectation value. The evolution of the expectation value $\langle X_1X_2X_3X_4 \rangle$ for deterministic pumping ($p = 1$) is also shown. The observed decay of $\langle Z_iZ_j \rangle$ is due to imperfections and is detrimental to the pumping process (see Supplementary Information). Error bars (**c** and **d**), $\pm 1\sigma$.

way, which is also an essential ingredient in quantum error correction protocols. Here, we illustrate this for a four-qubit stabilizer operator $X_1X_2X_3X_4$. As above, we first coherently map the information about whether the system spins are in the $+1$ (-1) eigenspace of the stabilizer operator onto the logical states $|0\rangle$ and $|1\rangle$ of the ancilla qubit. In contrast to the engineering of coherent and dissipative maps above, where this step was followed by single- and two-qubit gate operations, here we proceed instead by measuring the ancilla qubit.

Thus, depending on the measurement outcome for the ancilla, the system qubits are projected onto the corresponding eigenspace of the stabilizer: $\rho_S \mapsto P_{\pm} \rho_S P_{\pm} / N_{\pm}$ ($P_{\pm} = \rho_S P_{\pm} / N_{\pm}$) for finding the ancilla in $|0\rangle$ ($|1\rangle$) with the normalization factor $N_{\pm} = \text{Tr}(P_{\pm} \rho_S P_{\pm})$. Here, $P_{\pm} = \frac{1}{2}(1 \pm X_1X_2X_3X_4)$ denote the projectors onto the ± 1 eigenspaces of the stabilizer operator. Note that our measurement is QND in the sense that (superposition) states within one of the two eigenspaces are not affected by the measurement.

In the experiment with 4+1 ions, we prepare different four-qubit system input states (tomographically characterized in additional experiments), carry out the QND measurement and tomographically determine the resulting system output states.

To characterize how well the measurement device prepares a definite state, we use as input $|1111\rangle$, which is a non-eigenstate of the stabilizer. In this case, when the ancilla qubit is found in $|0\rangle$ or $|1\rangle$ the system qubits are prepared in the state $(|0000\rangle \pm |1111\rangle)/\sqrt{2}$ by the QND measurement. Experimentally we observe this behaviour with a quantum state preparation (QSP) fidelity⁴⁶ of $F_{\text{QSP}} = 73(1)\%$. On the other hand, for a stabilizer eigenstate, the QND measurement preserves the stabilizer expectation value. Experimentally, for the input state $(|0011\rangle - |1100\rangle)/\sqrt{2}$, we observe a QND fidelity⁴⁶ of $F_{\text{QND}} = 96.9(6)\%$. For more details, see Supplementary Information.

Conclusions and outlook

In the present work, we have demonstrated engineering of dissipative Kraus maps for Bell-state and four-qubit stabilizer pumping. These particular examples exploited the available quantum resources by

coupling the system qubits to an ancilla by a universal set of entangling operations. The engineered environment was here represented by an ancilla ion undergoing optical pumping by dissipative coupling to the vacuum modes of the radiation field. These experiments, where the ancilla remains unobserved, represent an open-loop dynamics. Such scenarios were recently discussed in the context of an open-system quantum simulator for spin models, including lattice gauge theories, realized with Rydberg atoms in optical lattices. In fact, four-qubit stabilizer pumping together with four-spin interactions demonstrate the basic ingredients²⁸ for the simulation of spin dynamics and ground-state cooling for Kitaev's toric code Hamiltonian³⁰ on a single four-spin plaquette.

For a closed system, only a small number of Hamiltonians as generators are required to generate all possible unitary time evolutions. In the context of qubits, this is given by a finite set of single qubit operations together with an entangling CNOT gate. In contrast, as noted in refs 25 and 29, when a single ancilla qubit is used, the most general Markovian open-system dynamics cannot be obtained with a finite set of non-unitary open-loop transformations. However, such a universal dynamical control can be achieved through repeated application of coherent control operations and measurement of the ancilla qubit, followed by classical feedback operations onto the system. We note that our demonstration of a multi-qubit QND measurement provides, in combination with our previously demonstrated feedback techniques⁴⁷, the basic ingredient for the realization of such closed-loop dynamics.

Our experimental demonstration of a toolbox of elementary building blocks in a system of trapped ions should be seen as a conceptual step towards the realization of an open quantum system simulator with applications in various fields¹⁴, including condensed-matter physics and quantum chemistry, possibly in modelling quantum effects in biology⁴⁸, and in quantum computation driven by dissipation²².

Although the present experiments were performed with a linear ion-trap quantum computer architecture, the continuing development of two-dimensional trap arrays promises scalable implementations of spin-model simulators. In addition, gate-based simulation approaches can incorporate quantum error correction protocols, which may prove essential for fault-tolerant quantum simulation. The demonstrated concepts can also be readily adapted to other physical platforms, ranging from optical, atomic and molecular systems to solid-state devices.

Received 12 October 2010; accepted 4 January 2011.

- Ladd, T. D. *et al.* Quantum computers. *Nature* **464**, 45–53 (2010).
- Kimble, H. J. The quantum internet. *Nature* **453**, 1023–1030 (2008).
- Schoelkopf, R. J. & Girvin, S. M. Wiring up quantum systems. *Nature* **451**, 664–669 (2008).
- Neeley, M. *et al.* Generation of three-qubit entangled states using superconducting phase qubits. *Nature* **467**, 570–573 (2010).
- Saffman, M., Walker, T. G. & Mølmer, K. Quantum information with Rydberg atoms. *Rev. Mod. Phys.* **82**, 2313–2363 (2010).
- Bloch, I., Dalibard, J. & Zwierger, W. Many-body physics with ultracold gases. *Rev. Mod. Phys.* **80**, 885–964 (2008).
- O'Brien, J. L. Optical quantum computing. *Science* **318**, 1567–1570 (2007).
- Jones, J. A. Quantum computing with NMR. Preprint at (<http://arXiv.org/abs/1011.1382>) (2010).
- Clarke, J. & Wilhelm, F. K. Superconducting quantum bits. *Nature* **453**, 1031–1042 (2008).
- Hanson, R., Kouwenhoven, L. P., Petta, J. R., Tarucha, S. & Vandersypen, L. M. K. Spins in few-electron quantum dots. *Rev. Mod. Phys.* **79**, 1217–1265 (2007).
- Wrachtrup, J. & Jelezko, F. Processing quantum information in diamond. *J. Phys. Condens. Matter* **18**, S807–S824 (2006).
- Feynman, R. Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982).
- Lloyd, S. Universal quantum simulators. *Science* **273**, 1073–1078 (1996).
- Buluta, I. & Nori, F. Quantum simulators. *Science* **326**, 108–111 (2009).
- Myatt, C. J. *et al.* Decoherence of quantum superpositions through coupling to engineered reservoirs. *Nature* **403**, 269–273 (2000).
- Viola, L. *et al.* Experimental realization of noiseless subsystems for quantum information processing. *Science* **293**, 2059–2063 (2001).
- Deléglise, S. *et al.* Reconstruction of nonclassical cavity field states with snapshots of their decoherence. *Nature* **455**, 510–514 (2008).

- Barreiro, J. T. *et al.* Experimental multiparticle entanglement dynamics induced by decoherence. *Nature Phys.* **6**, 943–946 (2010); published online 26 September 2010.
- Krauter, H. *et al.* Entanglement generated by dissipation. Preprint at (<http://arXiv.org/abs/1006.4344>) (2010).
- Diehl, S. *et al.* Quantum states and phases in driven open quantum systems with cold atoms. *Nature Phys.* **4**, 878–883 (2008).
- Cho, J., Bose, S. & Kim, M. S. Optical pumping into many-body entanglement. *Phys. Rev. Lett.* **106**, 020504 (2011).
- Verstraete, F., Wolf, M. M. & Cirac, J. I. Quantum computation and quantum-state engineering driven by dissipation. *Nature Phys.* **5**, 633–636 (2009).
- Pastawski, F., Clemente, L. & Cirac, J. I. Quantum memories based on engineered dissipation. Preprint at (<http://arXiv.org/abs/1010.2901>) (2010).
- Goldstein, G. *et al.* Environment-assisted precision measurement. Preprint at (<http://arXiv.org/abs/1001.0089>) (2010).
- Lloyd, S. & Viola, L. Engineering quantum dynamics. *Phys. Rev. A* **65**, 010101 (2001).
- Häffner, H., Roos, C. F. & Blatt, R. Quantum computing with trapped ions. *Phys. Rep.* **469**, 155–203 (2008).
- Home, J. P. *et al.* Complete methods set for scalable ion trap quantum information processing. *Science* **325**, 1227–1230 (2009).
- Weimer, H., Müller, M., Lesanovsky, I., Zoller, P. & Büchler, H. P. A Rydberg quantum simulator. *Nature Phys.* **6**, 382–388 (2010).
- Bacon, D. *et al.* Universal simulation of Markovian quantum dynamics. *Phys. Rev. A* **64**, 062302 (2001).
- Kitaev, A. Y. Fault-tolerant quantum computation by anyons. *Ann. Phys.* **303**, 2–30 (2003).
- Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).
- Steane, A. M. Efficient fault-tolerant quantum computing. *Nature* **399**, 124–126 (1999).
- Dür, W., Bremner, M. J. & Briegel, H. J. Quantum simulation of interacting high-dimensional systems: the influence of noise. *Phys. Rev. A* **78**, 052325 (2008).
- Wiseman, H. M. & Milburn, G. J. *Quantum Measurement and Control* (Cambridge Univ. Press, 2009).
- Schmidt-Kaler, F. *et al.* How to realize a universal quantum gate with trapped ions. *Appl. Phys. B* **77**, 789–796 (2003).
- Mølmer, K. & Sørensen, A. Multiparticle entanglement of hot trapped ions. *Phys. Rev. Lett.* **82**, 1835–1838 (1999).
- Monz, T. *et al.* Coherence of large-scale entanglement. Preprint at (<http://arXiv.org/abs/1009.6126>) (2010).
- Roos, C. F. *et al.* Control and measurement of three-qubit entangled states. *Science* **304**, 1478–1480 (2004).
- Gilchrist, A., Langford, N. K. & Nielsen, M. A. Distance measures to compare real and ideal quantum processes. *Phys. Rev. A* **71**, 062310 (2005).
- Jozsa, R. Fidelity for mixed quantum states. *J. Mod. Opt.* **41**, 2315–2323 (1994).
- Gühne, O. & Seevinck, M. Separability criteria for genuine multiparticle entanglement. *N. J. Phys.* **12**, 053002 (2010).
- Friedenauer, A., Schmitz, H., Glueckert, J., Porras, D. & Schaetz, T. Simulating a quantum magnet with trapped ions. *Nature Phys.* **4**, 757–761 (2008).
- Kim, K. *et al.* Quantum simulation of frustrated Ising spins with trapped ions. *Nature* **465**, 590–593 (2010).
- Porras, D. & Cirac, J. I. Effective quantum spin systems with trapped ions. *Phys. Rev. Lett.* **92**, 207901 (2004).
- Nayak, C., Simon, S. H., Stern, A., Freedman, M. & Das Sarma, S. Non-abelian anyons and topological quantum computation. *Rev. Mod. Phys.* **80**, 1083–1159 (2008).
- Ralph, T. C., Bartlett, S. D., O'Brien, J. L., Pryde, G. J. & Wiseman, H. M. Quantum nondemolition measurements for quantum information. *Phys. Rev. A* **73**, 012113 (2006).
- Riebe, M. *et al.* Deterministic entanglement swapping with an ion-trap quantum computer. *Nature Phys.* **4**, 839–842 (2008).
- Fleming, G. R., Huelga, S. & Plenio, M. (eds) *New J. Phys.* **12** (Focus on quantum effects and noise in biomolecules), (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Hammerer, I. Chuang, and O. Gühne for discussions and T. Northup for critically reading the manuscript. We acknowledge support by the Austrian Science Fund (FOQUS), the European Commission (AQUITE), the Institut für Quanteninformation GmbH, and a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme.

Author Contributions M.M. and J.T.B. developed the research, based on theoretical ideas proposed originally by P.Z.; J.T.B., P.S. and D.N. performed the experiments; J.T.B., P.S. and T.M. analysed the data; P.S., J.T.B., D.N., T.M., M.C., M.H. and R.B. contributed to the experimental set-up; M.M., J.T.B. and P.Z. wrote the manuscript, with revisions provided by C.F.R.; all authors contributed to the discussion of the results and manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.Z. (Peter.Zoller@uibk.ac.at; for theory) and R.B. (Rainer.Blatt@uibk.ac.at; for experiment).

Post-traumatic stress disorder is associated with PACAP and the PAC1 receptor

Kerry J. Ressler^{1,2,4}, Kristina B. Mercer¹, Bekh Bradley^{2,3}, Tanja Jovanovic², Amy Mahan⁴, Kimberly Kerley¹, Seth D. Norrholm^{2,3}, Varun Kilaru², Alicia K. Smith², Amanda J. Myers⁵, Manuel Ramirez⁵, Anzhelika Engel⁵, Sayamwong E. Hammack⁶, Donna Toufexis^{4,6}, Karen M. Braas⁷, Elisabeth B. Binder^{2,8} & Victor May⁷

Pituitary adenylate cyclase-activating polypeptide (PACAP) is known to broadly regulate the cellular stress response. In contrast, it is unclear if the PACAP–PAC1 receptor pathway has a role in human psychological stress responses, such as post-traumatic stress disorder (PTSD). Here we find, in heavily traumatized subjects, a sex-specific association of PACAP blood levels with fear physiology, PTSD diagnosis and symptoms in females. We examined 44 single nucleotide polymorphisms (SNPs) spanning the PACAP (encoded by *ADCYAPI*) and PAC1 (encoded by *ADCYAP1R1*) genes, demonstrating a sex-specific association with PTSD. A single SNP in a putative oestrogen response element within *ADCYAP1R1*, rs2267735, predicts PTSD diagnosis and symptoms in females only. This SNP also associates with fear discrimination and with *ADCYAP1R1* messenger RNA expression in human brain. Methylation of *ADCYAP1R1* in peripheral blood is also associated with PTSD. Complementing these human data, *ADCYAP1R1* mRNA is induced with fear conditioning or oestrogen replacement in rodent models. These data suggest that perturbations in the PACAP–PAC1 pathway are involved in abnormal stress responses underlying PTSD. These sex-specific effects may occur via oestrogen regulation of *ADCYAP1R1*. PACAP levels and *ADCYAP1R1* SNPs may serve as useful biomarkers to further our mechanistic understanding of PTSD.

PACAP was first isolated from ovine hypothalamic extracts based on its ability to stimulate cyclic AMP production in anterior pituitary cells¹. It is a highly conserved member of the VIP/secretin/glucagon peptide family, with pleiotropic functions in development, cell signalling, metabolism, homeostasis and cell protection^{2–6}. Among these myriad functions, studies have demonstrated (1) high expression of PACAP peptide and its selective PAC1 receptor in hypothalamic and limbic structures, (2) PACAP regulation of corticotropin releasing hormone and autonomic function, (3) actions of PACAP in stress-related behaviour, (4) reduced anxiety-like phenotypes in PACAP and PAC1 receptor null mice, and (5) blunted corticosterone response in knockout animals after emotional stressors. Thus, PACAP–PAC1 receptor signalling is integrally involved in stress mechanisms^{7,8}. We hypothesized that PACAPergic systems may be important mediators of abnormal stress responses following psychological trauma contributing to PTSD, which is an extreme maladaptive and debilitating psychiatric disorder affecting up to 40% of individuals over lifetime exposure to traumatic events^{9,10}.

Little is known about the biological processes regulating PTSD and other stress-related responses. To examine whether the PACAP–PAC1 pathway is associated with PTSD in a high risk, heavily traumatized population, we analysed blood levels of PACAP, and genetic variation and methylation of the PACAP (*ADCYAPI*) and PAC1 receptor (*ADCYAP1R1*) genes, in a cohort of more than 1,200 highly traumatized subjects with and without PTSD (see Supplementary Tables 1 and 2 for demographic information).

PACAP levels associated with PTSD in females

Using radioimmunoassay, we first examined PACAP peptide levels in peripheral blood samples from a previously described, highly traumatized, at risk population^{11–13} that had been matched on age, sex, and trauma histories ($n = 64$, see Supplementary Tables 1–3 for demographics). We found that PTSD symptoms (PTSD symptom scale¹⁴) were significantly correlated with PACAP38 (PACAP peptide containing 38 residues) blood levels in females ($P < 0.005$, $r = 0.497$, Fig. 1a), but not in males ($P > 0.5$). Also in females, PTSD diagnosis was associated with PACAP38 levels ($P \leq 0.001$), with higher PACAP38 found in the PTSD cohort. Furthermore, PACAP levels (median split, low versus high) were differentially associated with PTSD symptoms in females (Fig. 1b). PACAP38 levels also predicted differential response on all three symptom clusters necessary to fulfil diagnostic criteria for PTSD (intrusive re-experiencing (for example, trauma flashbacks), avoidance (for example, avoidance of trauma reminders) and hyperarousal (for example, increased startle response)) in females but not males (Fig. 1c). These analyses were repeated in a second, all female cohort ($N = 74$) with similar findings (Fig. 1d; high versus low PACAP38 levels, controlling for age, substance abuse and total trauma exposure, one-tailed t -tests: total symptoms, $P \leq 0.05$, hyperarousal symptoms, $P \leq 0.001$; and percentage with clinically significant symptoms, $\chi^2 = 4.9$, $P < 0.05$). These observations were especially notable, as females may be at twice the risk for PTSD as compared to males^{9,11}, implicating roles for sex hormones, especially oestrogen, in the disorder^{15–17}. When we controlled for

¹Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA. ²Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, Atlanta, Georgia 30322, USA. ³Atlanta VA Medical Center, Atlanta, Georgia 30033, USA. ⁴Yerkes National Primate Research Center, Atlanta, Georgia 30329, USA. ⁵University of Miami, Miller School of Medicine, Miami, Florida 33136, USA. ⁶Department of Psychology, University of Vermont, Burlington, Vermont 05401, USA. ⁷Departments of Anatomy and Neurobiology and Pharmacology, University of Vermont College of Medicine, Burlington, Vermont 05401, USA. ⁸Max Planck Institute of Psychiatry, Munich 80804, Germany.

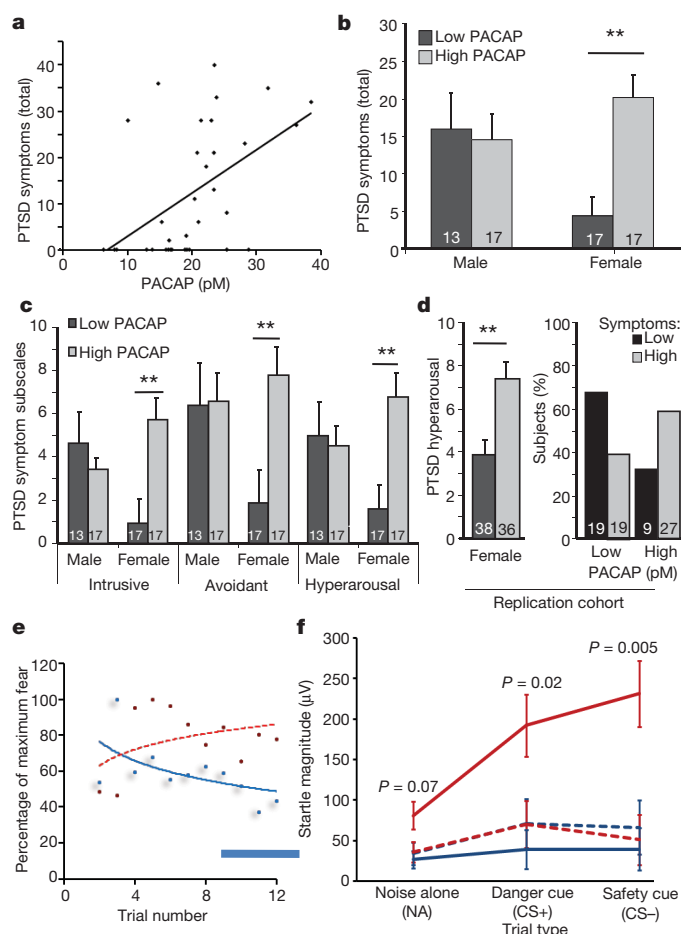


Figure 1 | PACAP blood levels predict PTSD symptoms in females. **a**, PTSD symptoms (scale range 0–51), relative to plasma PACAP38 blood levels (pM); ($N = 34$ females; $r = 0.497$, $P \leq 0.005$). **b**, Total PTSD symptoms plotted relative to sex and levels of plasma PACAP38 ($N = 64$, low: <20 pM, high: >20 pM); females with high PACAP blood levels have increased symptoms (** $P < 0.0005$). **c**, PACAP levels (low versus high) are also differentially associated with PTSD intrusive, avoidance and hyperarousal symptoms in females ($N = 64$, ** $P < 0.005$). **d**, PACAP levels (low versus high) were examined in a replication sample of highly traumatized women, with differential association in hyperarousal symptoms (left, $N = 74$, ** $P = 0.002$) and in the percentage of subjects with significant symptoms (right, $\chi^2 = 4.9$, $P < 0.05$). **e**, Acoustic startle reflex (EMG) relative to the fear conditioning trial in subjects without PTSD (blue) versus with PTSD (red). Habituation is seen in non-PTSD subjects during late acquisition (bar). **f**, Startle magnitude during the late acquisition period versus trial type (noise alone, CS+ and CS–), showing that females with high PACAP levels show enhanced startle responses to both fear cues (CS+, $P = 0.02$) and safety cues (CS–, $P = 0.005$) ($N = 27$; 16 male, 11 female). Dashed lines, low PACAP; solid lines, high PACAP; blue, male; red, female. Bars represent mean \pm s.e.m., N values for each group at bottom of bar graphs.

common stress-related phenomena (depression and history of substance abuse), the effect of PACAP level on PTSD remained ($P < 0.05$). In contrast, there was no effect of PACAP level on depression symptoms or substance abuse when controlling for PTSD.

In addition to the psychological symptoms that define the syndrome, subjects with PTSD have been found to have abnormally high conditioned fear responses. This high level of fear may result from a combination of an inability to habituate to aversive stimuli, a decreased ability to extinguish (learn to inhibit) fear memories, and possibly an over-consolidation of the original fear memory^{18–22}. Hence, we determined the physiological (electromyographic) levels of conditioned fear for 27 participants (16 male, 11 female) with PACAP blood levels. Fear potentiation was determined by measuring

the acoustic startle reflex in the presence of startle cues alone, or startle cues combined with stimuli paired (conditioned stimulus, CS+) or unpaired (CS–) with an aversive airblast. Female (but not male) subjects with high PACAP38 levels demonstrated markedly increased startle reflex responses to both CS+ ($P = 0.02$) and CS– ($P = 0.005$) cues. This was particularly pronounced during the late acquisition phase when normal subjects had habituated to the fearful stimuli (Fig. 1e, f). In aggregate, these data suggest that PACAP38 peptide is strongly associated with the psychological and physiological symptoms of PTSD in women with a history of trauma.

ADCYAP1R1 associated with PTSD in females

To assess whether there may be a genetic association of PTSD with polymorphisms in either the PACAP (*ADCYAP1*) or PAC1 receptor (*ADCYAP1R1*) locus, we performed a tag-SNP analysis ($r^2 = 0.8$; minor allele frequency (MAF) = 0.1) across both genes with a total of 44 SNPs (14 *ADCYAP1* and 30 *ADCYAP1R1* SNPs). Using logistic regression, we examined whether each SNP was associated with PTSD diagnosis in this cohort of highly traumatized urban civilian subjects ($n = 798$)^{11,12,23}, in total, or stratified by gender (females: $n = 503$; males: $n = 295$). Only the *ADCYAP1R1* receptor SNP rs2267735 ($P = 0.0002$ in females; NS in males) remained significant after experiment-wide multiple correction for sex and 44 independent tests (Fig. 2a and b, Supplementary Fig. 1). No SNPs in the peptide *ADCYAP1* gene met experiment-wide criteria for association (Supplementary Fig. 2). Given these striking gender differences and recent data demonstrating that *ADCYAP1R1* gene expression may be dynamically modulated by oestrogen²⁴, the distribution of oestrogen response elements (EREs) within the *ADCYAP1R1* gene was examined (Supplementary Table 4). We found that rs2267735 was within a predicted ERE (Fig. 2c, Genomatix; matrix similarity = 0.877, core similarity = 1.0). Because rs2267735 is positioned within the central variable region of the consensus sequence, *in silico* analyses do not currently allow us to predict how the ‘C’ versus ‘G’ allele may differentially alter ERE function, and further *in vitro* analyses are warranted.

We next determined if the association between rs2267735 and PTSD diagnosis could be replicated in an additional 439 subjects. These subjects were from the same overall study, but were interviewed and had DNA collected after the original discovery population. Thus they served as a replication source from the same population but distinct in time and with different interviewing staff. The table in Fig. 3a shows the logistic regression results for males and females separately in the initial population described in the tag-SNP analysis, the replication sample from the same population, and the combined sample of 1,237 individuals. The main effect of the SNP on PTSD diagnosis could be replicated in women ($P < 0.05$) and combining both samples increased the significance of the association ($N = 763$, $P < 0.00002$). As in the discovery sample, no effects were observed in males (male combined sample $N = 474$, $P = 0.7$).

To further examine *ADCYAP1R1* rs2267735 SNP associations with continuous PTSD symptom levels in females, we analysed both an additive and a dominant model with total PTSD symptoms and symptom subscales using the combined samples (Fig. 3b–e). The ‘CC’ genotype was most robustly associated with total PTSD symptoms, and among subscales, hyperarousal symptoms were the most strongly associated with rs2267735. Notably, even after controlling for childhood trauma history and adult trauma, age and race (which slightly reduces total N owing to missing data), the rs2267735 ‘CC’ genotype was associated with higher levels of PTSD hyperarousal symptoms compared to ‘G’ carriers in women ($P = 0.0008$, Fig. 3e), but not men ($P = 0.51$).

We repeated the above analyses with Beck depression inventory (BDI) symptoms and history of life-time substance abuse, and found no associations with these measures and rs2267735 (Supplementary Fig. 3), suggesting that this association may be relatively specific to PTSD. To address whether rs2267735 might be associated with other severe psychiatric illnesses, we performed analyses using bipolar

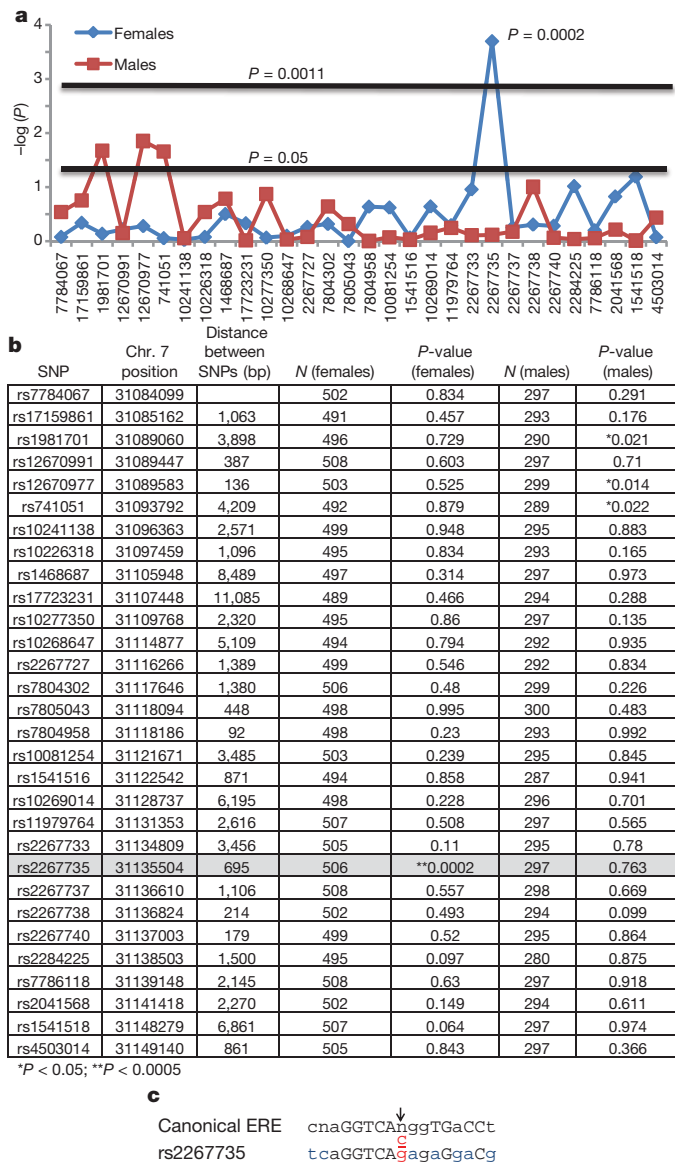


Figure 2 | Genetic association of PAC1 receptor (*ADCYAP1R*) with PTSD. **a**, 30 SNPs spanning the *ADCYAP1R* gene (x axis), with the $-\log(P)$ -value of logistic regression for each SNP association with PTSD (diagnosis based on DSM-IV criteria from PTSD Symptom Scale). Subjects were analysed with logistic regression in females only ($N = 503$) or males only ($N = 295$). Horizontal lines represent the nominal $P = 0.05$ or the corrected P -value, $P = 0.0011$ (44 SNPs, correcting for 30 *ADCYAP1R* SNPs and 14 *ADCYAP1* SNPs; Supplementary Fig. 1). rs2267735 is the only SNP remaining significant after multiple corrections ($P = 0.0002$). **b**, Table of P -values resulting from the association of each genotyped, *ADCYAP1R* SNP with PTSD diagnosis (by gender). The location on chromosome 7 for each SNP including the distance (bp) between the SNPs is given. The average distance between SNPs is 2.5 kb. SNP rs2267735 is located in an intron of *ADCYAP1R*, and is not in linkage disequilibrium with other SNPs (for African Americans in our population, data not shown). **c**, rs2267735 (C/G), in red, is located within a canonical oestrogen response element (ERE) binding site (capital letters, conserved canonical ERE nucleotides; blue letters, mismatches with the *ADCYAP1R* gene and canonical ERE; reverse strand shown).

disorder, schizophrenia, and Alzheimer's disease samples. From the data of the Genetic Association Information Network (GAIN) publicly accessible database (<http://www.genome.gov/19518664>), we analysed the association of rs2267735 (included on the Affymetrix 6.0 SNP array) with bipolar disorder as well as schizophrenia. We did not observe a significant association of this SNP with these two disorders in subjects with African American (954 cases, 1,195 controls) or

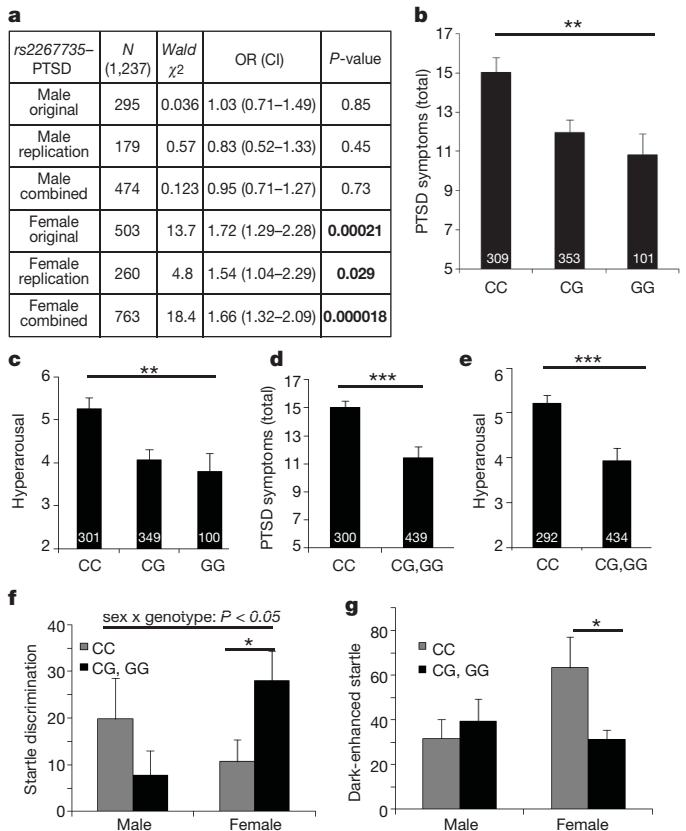


Figure 3 | Association of *ADCYAP1R* rs2267735 with PTSD symptoms and physiological fear responses. **a**, Table demonstrating the N , Wald χ^2 , odds ratio (OR) for C as risk allele and P -value, in males and females, in the original, replication and combined samples for logistic regression of rs2267735 with PTSD diagnosis. CI, confidence interval. **b**, Total PTSD symptoms are differentially associated with rs2267735 genotype in females ($P \leq 0.001$). **c**, Hyperarousal is the most robustly associated symptom with rs2267735 genotype ($P = 0.0009$). **d**, In a dominant/recessive model, even after controlling for childhood trauma, adult trauma and age, genotype predicts total PTSD symptoms ($P \leq 0.001$) and **e**, hyperarousal symptoms ($P \leq 0.0001$). **f**, Fear discrimination, measured with potentiated startle (CS+ startle minus CS– startle) is impaired in females with rs2267735 'CC' genotype. **g**, Dark-enhanced startle (startle^{dark} – startle^{light}) is significantly increased in females with rs2267735 'CC' genotype. N values are shown at base of each bar, bars represent mean \pm s.e.m. N values are slightly different across analyses owing to differences in number of subjects across measures. * $P < 0.05$; ** $P < 0.001$; *** $P < 0.0002$.

European (1,378 cases, 1,351 controls) ancestry. Specifically, we found that all pre-computed P -values for associations of rs2267735 with schizophrenia or bipolar disorder were higher than P uncorrected = 0.01, indicating no major contribution of this variant.

Additionally, we examined the association of rs2267735 and Alzheimer's in a previously characterized Alzheimer's disease sample²⁵. In this cohort of 342 subjects, we found no association with rs226735 and Alzheimer's disease diagnosis using either the additive genetic model ($P = 0.19$) or the dominant/recessive model ($P = 0.89$). These data suggest that we find robust associations with rs2267735 in women, but not men, with PTSD. In contrast, we find no association with depression symptoms, substance abuse, Alzheimer's disease, bipolar disorder, or schizophrenia across different samples. Note that for all of these negative results, owing to the limited sample sizes, we cannot rule out the possibility that rs2267735 may be associated with PTSD in men or with other disorders with a smaller effect size than we see with PTSD in women.

To parallel our results with plasma PACAP38 levels, we next examined whether physiological measures of fear are differentially associated with the *ADCYAP1R* rs2267735 SNP. In PTSD, but not depression¹⁸, fear response to an inhibitory CS–, or 'safety signal', is

exaggerated. The discrimination between CS+ and CS− improves across the training procedure in controls, but not in those with PTSD. We examined whether rs2267735 was associated with impaired fear discrimination late in conditioned acquisition, during the same period noted in Fig. 1e. Notably, females with the ‘CC’ genotype were significantly less able to discriminate CS+ from CS− signals (Fig. 3f, sex × genotype interaction, $P < 0.05$, and ‘CC’ versus ‘G’ carriers in females, $P < 0.05$).

We next examined whether a difference in dark-enhanced startle, a measure of increased anxiety in humans that is similar to light-enhanced startle in rodents^{26–28}, was differentially associated with rs2267735. Again, we found that females, but not males, with the ‘CC’ genotype showed significantly more startle in the dark compared to the light (Fig. 3g, males, $N = 35$, $P = 0.71$; females, $N = 53$, $P = 0.02$). Together, these data suggest that the *ADCYAP1R1* rs2267735 SNP may be relatively specific in its association with PTSD psychological and physiological phenotypes. Further, the robust association of rs2267735 with hyperarousal symptoms suggests that the role of PACAP–PAC1 may be specifically involved in the normalization of the stress response, a process which is particularly dysregulated in PTSD.

ADCYAP1R1 methylation and mRNA expression

Environmental, genetic and epigenetic mechanisms are likely to moderate the long-term effects of trauma exposure. Using the Illumina HumanMethylation27 BeadChip, we interrogated methylation in DNA extracted from peripheral blood at the first site within the *ADCYAP1R1* CpG island (Supplementary Fig. 2). Methylation at this site was significantly associated with total PTSD symptoms (Fig. 4a, $N = 94$, $r = 0.354$, $P < 0.0005$) in a sex-independent manner. Further, CpG methylation level (median split) was associated with PTSD

diagnosis (Fig. 4b, $\chi^2 = 8.1$, $P < 0.005$), but not depression ($P > 0.05$, Supplementary Fig. 3e). There was no significant association between methylation of *ADCYAP1R1* and PTSD symptoms. These data suggest that *ADCYAP1R1* is regulated, in part, through epigenetic mechanisms that contribute to differential function of the PAC1 receptor in PTSD.

To examine the potential relationship of genotype and brain mRNA expression as described previously²⁹, we used a brain mRNA expression data set³⁰ to test whether *ADCYAP1R1* rs2267735 is associated with differential gene expression. We first examined whether cortical *ADCYAP1R1* and *ADCYAP1* mRNA levels were correlated. As shown in Fig. 4c, these mRNA levels were significantly inversely correlated ($r = -0.219$, $P < 0.001$, including males and females), suggesting that brain levels of PACAP peptide and PAC1 mRNA are tightly regulated.

We next used a previously analysed data set with combined genome-wide association and brain mRNA expression data³⁰ to examine whether *ADCYAP1R1* rs2267735 imputed genotypes were associated with differential *ADCYAP1R1* expression in brain. We found a sex × genotype effect (Fig. 4d, $F(3,99) = 4.3$, $P < 0.05$) with females with the ‘CC’ genotype expressing significantly less *ADCYAP1R1* mRNA than males ($F(1,33) = 5.5$, $P < 0.05$) or than females who are ‘G’ carriers (one-tailed, $F(1, 45) = 2.87$, $P < 0.05$). Thus, mRNA encoding the PACAP peptide and PAC1 receptor appeared to be tightly regulated within the human cortex, and *ADCYAP1R1* mRNA levels were associated with the *ADCYAP1R1* rs2267735 SNP.

Fear induces *Adcyap1r1* in mouse amygdala

Despite prior studies examining PACAP–PAC1 receptor function in central/peripheral nervous system development, endocrine homeostasis, metabolism, cellular protection/regeneration and chronic stress responses^{2,3,6,31–34}, a role for PACAP signalling in fear conditioning has not been evaluated. Given our data implicating PACAP in PTSD, we wondered if *Adcyap1r1* mRNA was differentially regulated in mice using Pavlovian fear conditioning^{35–40}, a means of studying acute fear and trauma responses that has been proposed to model PTSD^{19,22}. We performed classical fear conditioning experiments using male mice, in which a previously neutral tone CS (6 kHz) was paired with 10 foot-shocks (1 mA, 0.5 s; Fig. 5a). This conditioning paradigm consistently provides robust fear learning in mice leading to changes in gene expression within the amygdala, a region critical for fear learning and expression. Quantitative PCR analyses shows that amygdala *Adcyap1r1* mRNA increased ~1.5-fold during the consolidation of fear (Fig. 5b, $P < 0.05$), with a similar trend within the medial prefrontal cortex (mPFC). When peak freezing was compared with brain mRNA levels, we find a significant correlation between fear learning and *Adcyap1r1* mRNA (Fig. 5c, $r^2 = 0.49$, $P < 0.05$).

Oestrogen induces *Adcyap1r1* in rat BNST

To further establish the relationship between PACAP–PAC1 receptors and oestrogen in a validated model of sex hormone regulation, we examined oestrogen-induced changes in *Adcyap1* and *Adcyap1r1* transcripts in the bed nucleus of stria terminalis (BNST) in female rats. The BNST is a component of the extended amygdala that is subject to significant gonadal hormonal control^{7,27,28}. In rodents, it is critical for emotional behaviour, mediating stress responses and the light-enhanced startle response. We examined gene expression in the BNST in ovariectomized female rats following 21-day implantation of continuous release oestrogen pellets. Compared to control implants, oestradiol increased *Adcyap1* transcripts in the dorsal and ventral BNST 2.1- and 3.4-fold, respectively ($P \leq 0.01$, Fig. 5d). Additionally, oestradiol increased *Adcyap1r1* mRNA 1.5-fold in the dorsal BNST samples ($P < 0.05$, Fig. 5e), and future studies should also examine oestradiol sensitivity of these genes in amygdala. While these rodent studies are complex and have differing experimental designs, our data clearly illustrate dynamic PACAP–PAC1 receptor regulation within central areas mediating fear, stress and oestrogen responsiveness.

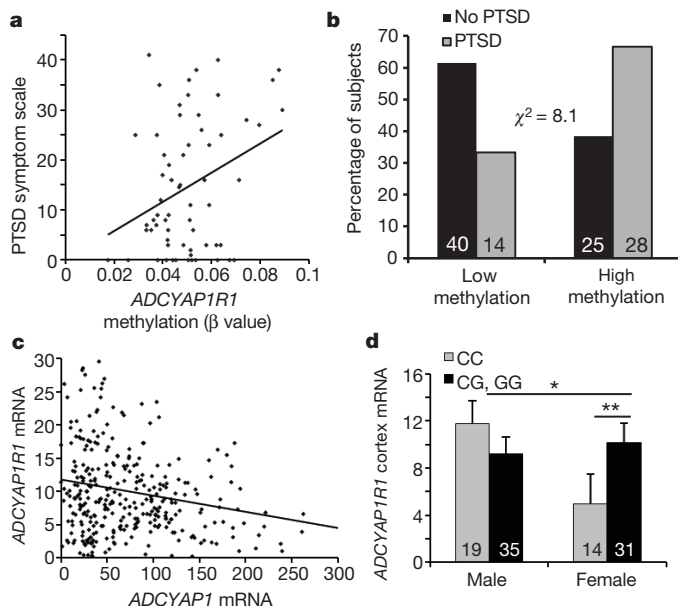


Figure 4 | *ADCYAP1R1* methylation and mRNA expression. **a**, Methylation within the first CpG island of *ADCYAP1R1* (β value, Illumina cg27076139) is positively correlated with total PTSD symptoms (both sexes; $N = 107$; $r = 0.354$, $P < 0.0005$). **b**, Subjects with PTSD have higher levels of *ADCYAP1R1* methylation (median split, $N = 107$; χ^2 analyses, $P < 0.005$). **c**, *ADCYAP1* mRNA levels are inversely correlated with *ADCYAP1R1* mRNA levels in cortex (from prior data set¹³) ($r = -0.219$; $P < 0.001$). **d**, *ADCYAP1R1* mRNA levels are differentially expressed in females compared to males based on imputed *ADCYAP1R1* rs2267735 genotype (from prior data set¹³) (* $P < 0.05$ male versus female CC carriers, ** $P < 0.05$, one-tailed, CC versus G-carriers). Bars represent mean \pm s.e.m. N values for each group at bottom of each graph.

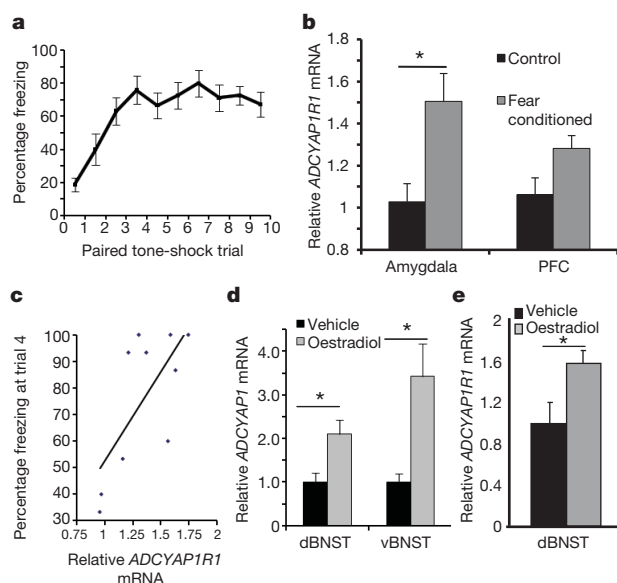


Figure 5 | Regulation of *Adcyap1r1* and *Adcyap1* mRNA in rodent models. **a**, Percentage of time freezing, in mice, to the conditioned tone (CS+) following tone-shock pairings during the conditioned fear trials. **b**, RT-PCR analyses of mRNA levels within mouse amygdala and mPFC 2 h after fear conditioning or in control handling conditions, showing a significant increase in amygdala *Adcyap1r1* mRNA ($N = 15$, 1.47 fold, $P < 0.05$) and a non-significant trend in mPFC (1.19 fold change). **c**, Correlation between average amygdala and PFC *Adcyap1r1* mRNA and percentage freezing at trial 4, demonstrating an association between *Adcyap1r1* mRNA with rate of fear learning ($r^2 = 0.49$, $P < 0.05$). **d**, *Adcyap1* mRNA in rat BNST in female rats ($N = 12$ per group) following ovariectomy and oestradiol implant versus vehicle replacements. *Adcyap1* mRNA is increased in both dorsal BNST (dBNST, 2.1-fold) and ventral BNST (vBNST, 3.4-fold) after oestradiol implantation. **e**, *Adcyap1r1* transcripts are also increased in dorsal BNST (1.6-fold, $N = 4$ per group). Bars represent mean \pm s.e.m. * $P < 0.05$.

Discussion

Since its identification more than 20 years ago, PACAP has been increasingly implicated in diverse cellular stress response pathways and neurotrophic function. However, the organizational role of the PACAP system in orchestrating behavioural stress responses has yet to be clarified. Our data suggest that PACAP–PAC1 receptor expression and signalling may be integrally involved in regulating the psychological and physiological responses to traumatic stress. Further, we report an association of an ERE-embedded *ADCYAP1R1* SNP with PTSD, and we demonstrate fear- and oestrogen-dependent regulation of PACAP systems within stress-responsive regions of the brain. These data may begin to explain sex-specific differences in PTSD diagnosis, symptoms and fear physiology. Future work targeting the PACAP–PAC1 receptor system may lead to novel and robust biomarkers; it may also further our understanding of the neural mechanisms underlying pathological responses to stress, and help identify potential therapeutic targets for the prevalent and debilitating syndrome of PTSD.

METHODS SUMMARY

This highly traumatized, civilian, cross-sectional cohort has been previously described in candidate gene-association studies of PTSD and depression^{11–13}. Research interviews, salivary DNA and blood samples were collected from patients receiving services in the primary care clinics at Grady Memorial Hospital (Atlanta, Georgia, USA). All study procedures have been reviewed and approved by the Emory Institutional Review Board and the Grady Hospital Research Oversight Committee. PTSD measures in this manuscript are based on the PTSD symptom scale¹⁴, which has been validated within this population using the Clinician Administered PTSD Scale. PACAP38 radioimmunoassay (1:30,000, Peninsula Laboratories) was performed at University of Vermont, using double antibody immunoprecipitation as described⁴¹. For genotyping, pairwise tagging ($r^2 > 0.8$,

MAF > 0.1) was used to choose tag-SNPs for both *ADCYAP1* and *ADCYAP1R1*. The coordinates were chr. 18 885000–906000 and chr. 7 31048667–31117836 for *ADCYAP1* and *ADCYAP1R1*, respectively (NCBI B36), which includes approximately 10 kilobases (kb) upstream and 5 kb downstream of the coding regions for both genes. Genotypes for the tag-SNPs were generated using Sequenom iPLEX with follow-up analyses using Taqman. For methylation analyses, bisulphite-converted DNA was whole-genome amplified, fragmented, and hybridized to the HumanMethylation27 BeadChip (Illumina). Individual samples were stratified to separate BeadChips according to PTSD status to limit bias. The BeadChips were scanned using a BeadStation 500GX, and the methylation level (β value) was calculated using the Methylation Module of the BeadStudio software. The eyeblink component of the acoustic startle response was measured by electromyographic recordings of the right orbicularis oculi muscle with two 5-mm Ag/AgCl electrodes filled with electrolyte gel, as described^{18,19}. The mouse fear conditioning and rat oestrogen replacement studies are described in detail in Supplementary Methods.

Received 16 August 2010; accepted 6 January 2011.

- Miyata, A. *et al.* Isolation of a novel 38 residue-hypothalamic polypeptide which stimulates adenylate cyclase in pituitary cells. *Biochem. Biophys. Res. Commun.* **164**, 567–574 (1989).
- Ghzi, H. *et al.* Role of PACAP in the physiology and pathology of the sympathoadrenal system. *Front. Neuroendocrinol.* **29**, 128–141 (2008).
- Hashimoto, H., Shintani, N. & Baba, A. New insights into the central PACAPergic system from the phenotypes in PACAP- and PACAP receptor-knockout mice. *Ann. NY Acad. Sci.* **1070**, 75–89 (2006).
- Spengler, D. *et al.* Differential signal transduction by five splice variants of the PACAP receptor. *Nature* **365**, 170–175 (1993).
- Watanabe, J. *et al.* Localization, characterization and function of pituitary adenylate cyclase-activating polypeptide during brain development. *Peptides* **28**, 1713–1719 (2007).
- Zhong, Y. Mediation of PACAP-like neuropeptide transmission by coactivation of Ras/Raf and cAMP signal transduction pathways in *Drosophila*. *Nature* **375**, 588–592 (1995).
- Hammack, S. E. *et al.* Chronic stress increases pituitary adenylate cyclase-activating peptide (PACAP) and brain-derived neurotrophic factor (BDNF) mRNA expression in the bed nucleus of the stria terminalis (BNST): roles for PACAP in anxiety-like behavior. *Psychoneuroendocrinology* **34**, 833–843 (2009).
- Vaudry, D. *et al.* Pituitary adenylate cyclase-activating polypeptide and its receptors: 20 years after the discovery. *Pharmacol. Rev.* **61**, 283–357 (2009).
- Breslau, N. The epidemiology of posttraumatic stress disorder: what is the extent of the problem? *J. Clin. Psychiatry* **62** (Suppl 17), 16–22 (2001).
- Hoge, C. W., Auchterlonie, J. L. & Milliken, C. S. Mental health problems, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. *J. Am. Med. Assoc.* **295**, 1023–1032 (2006).
- Binder, E. B. *et al.* Association of FKBP5 polymorphisms and childhood abuse with risk of posttraumatic stress disorder symptoms in adults. *J. Am. Med. Assoc.* **299**, 1291–1305 (2008).
- Bradley, R. G. *et al.* Influence of child abuse on adult depression: moderation by the corticotropin-releasing hormone receptor gene. *Arch. Gen. Psychiatry* **65**, 190–200 (2008).
- Gillespie, C. F., Phifer, J., Bradley, B. & Ressler, K. J. Risk and resilience: genetic and environmental influences on development of the stress response. *Depress. Anxiety* **26**, 984–992 (2009).
- Foa, E. B. & Tolin, D. F. Comparison of the PTSD symptom scale-interview version and the clinician-administered PTSD scale. *J. Trauma. Stress* **13**, 181–191 (2000).
- Bangasser, D. A. *et al.* Sex differences in corticotropin-releasing factor receptor signaling and trafficking: potential role in female vulnerability to stress-related psychopathology. *Mol. Psychiatry* **15**, 896–904 (2010).
- McEwen, B. S. Steroid hormones: effect on brain development and function. *Horm. Res.* **37** (Suppl 3), 1–10 (1992).
- Shansky, R. M. *et al.* Estrogen promotes stress sensitivity in a prefrontal cortex-amygdala pathway. *Cereb. Cortex* **20**, 2560–2567 (2010).
- Jovanovic, T. *et al.* Impaired fear inhibition is a biomarker of PTSD but not depression. *Depress. Anxiety* **27**, 244–251 (2010).
- Jovanovic, T. & Ressler, K. J. How the neurocircuitry and genetics of fear inhibition may inform our understanding of PTSD. *Am. J. Psychiatry* **167**, 648–662 (2010).
- Rauch, S. L., Shin, L. M. & Phelps, E. A. Neurocircuitry models of posttraumatic stress disorder and extinction: human neuroimaging research—past, present, and future. *Biol. Psychiatry* **60**, 376–382 (2006).
- Shin, L. M. & Liberzon, I. The neurocircuitry of fear, stress, and anxiety disorders. *Neuropsychopharmacology* **35**, 169–191 (2010).
- Yehuda, R. & LeDoux, J. Response variation following trauma: a translational neuroscience approach to understanding PTSD. *Neuron* **56**, 19–32 (2007).
- Gillespie, C. F. *et al.* Trauma exposure and stress-related disorders in inner city primary care patients. *Gen. Hosp. Psychiatry* **31**, 505–514 (2009).
- Aenlle, K. K., Kumar, A., Cui, L., Jackson, T. C. & Foster, T. C. Estrogen effects on cognition and hippocampal transcription in middle-aged mice. *Neurobiol. Aging* **30**, 932–945 (2009).
- Corneveaux, J. J. *et al.* Association of CR1, CLU and PICALM with Alzheimer's disease in a cohort of clinically characterized and neuropathologically verified individuals. *Hum. Mol. Genet.* **19**, 3295–3301 (2010).

26. Grillon, C., Morgan, C. A. III, Davis, M. & Southwick, S. M. Effect of darkness on acoustic startle in Vietnam veterans with PTSD. *Am. J. Psychiatry* **155**, 812–817 (1998).
27. Walker, D. L. & Davis, M. Double dissociation between the involvement of the bed nucleus of the stria terminalis and the central nucleus of the amygdala in startle increases produced by conditioned versus unconditioned fear. *J. Neurosci.* **17**, 9375–9383 (1997).
28. Walker, D. L. & Davis, M. Light-enhanced startle: further pharmacological and behavioral characterization. *Psychopharmacology (Berl.)* **159**, 304–310 (2002).
29. McMahon, F. J. *et al.* Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1. *Nature Genet.* **42**, 128–131 (2010).
30. Myers, A. J. *et al.* A survey of genetic human cortical gene expression. *Nature Genet.* **39**, 1494–1499 (2007).
31. Botia, B. *et al.* Neurotrophic effects of PACAP in the cerebellar cortex. *Peptides* **28**, 1746–1752 (2007).
32. Brennum, D. E. Neuroprotection: a comparative view of vasoactive intestinal peptide and pituitary adenylate cyclase-activating polypeptide. *Peptides* **28**, 1720–1726 (2007).
33. Dejda, A. *et al.* Inhibitory effect of PACAP on caspase activity in neuronal apoptosis: a better understanding towards therapeutic applications in neurodegenerative diseases. *J. Mol. Neurosci.* **36**, 26–37 (2008).
34. Hamelink, C. *et al.* Pituitary adenylate cyclase-activating polypeptide is a sympathoadrenal neurotransmitter involved in catecholamine regulation and glucohomeostasis. *Proc. Natl Acad. Sci. USA* **99**, 461–466 (2002).
35. Davis, M. The role of the amygdala in fear and anxiety. *Annu. Rev. Neurosci.* **15**, 353–375 (1992).
36. Ehrlich, I. *et al.* Amygdala inhibitory circuits and the control of fear memory. *Neuron* **62**, 757–771 (2009).
37. Fanselow, M. S. & LeDoux, J. E. Why we think plasticity underlying Pavlovian fear conditioning occurs in the basolateral amygdala. *Neuron* **23**, 229–232 (1999).
38. LeDoux, J. E. Emotional memory: in search of systems and synapses. *Ann. NY Acad. Sci.* **702**, 149–157 (1993).
39. Maren, S. & Quirk, G. J. Neuronal signalling of fear memory. *Nature Rev. Neurosci.* **5**, 844–852 (2004).
40. Pape, H. C. & Pare, D. Plastic synaptic networks of the amygdala for the acquisition, expression, and extinction of conditioned fear. *Physiol. Rev.* **90**, 419–463 (2010).
41. Girard, B. A. *et al.* Noncompensation in peptide/receptor gene expression and distinct behavioral phenotypes in VIP- and PACAP-deficient mice. *J. Neurochem.* **99**, 499–513 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was primarily supported by NIH grant MH071537 (K.J.R.), as well as DA019624 (K.J.R.) and HD27468 (V.M.). Support was also received from Emory and Grady Memorial Hospital General Clinical Research Center, NIH National Centers for Research Resources (M01RR00039 and P20RR16435), the American Foundation for Suicide Prevention (B.B.) and the Burroughs Wellcome Fund (K.J.R.). RNA samples were run using the Oncogenomics Core Facility at University of Miami. A.J.M. is supported by The National Institute on Aging (AG034504). We thank J.F. Cubells, Y. Tang and K. Conneely for discussions. We thank the Grady Trauma Project, including C.F. Gillespie, A. Schwartz, A. Wingo, D.A. Gutman and T. Weiss for medical support; A. Graham, A. Brown, J. Phifer, D. Crain, A. Kamkwala, J. Poole, D. Cross, N. Fani and A. Smith for clinical research support; and K. Schutz, E. Reiser and C. Fitzgerald for molecular/genetics technical support. Methylation chip assays were performed by the Emory University Biomarkers Service Center.

Author Contributions K.J.R. designed the initial experiments and wrote the initial version of the paper. K.J.R., B.B. and V.M. organized collaborations, obtained funding, supervised data collection and analyses, and revised the paper. K.B.M., K.K. and E.B.B. performed the genetics experiments and analyses on the primary, replication and GAIN cohorts, and revised the paper. T.J. and S.D.N. performed and supervised the human physiology studies and revised the paper. V.K. and A.K.S. performed and supervised the methylation studies and revised the paper. A.M. and K.J.R. performed and supervised the rodent fear conditioning and related RT-PCR analysis. D.T., S.E.H. and V.M. performed and supervised the oestrogen replacement studies and related RT-PCR analyses, and assisted with paper revisions. A.J.M., M.R. and A.E. performed and supervised the human mRNA expression analyses and the Alzheimer's disease gene association studies. K.M.B. and V.M. performed and supervised the PACAP38 radioimmunoassays and related data analyses.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.J.R. (kressle@emory.edu).

Transient Hoogsteen base pairs in canonical duplex DNA

Evgenia N. Nikolova¹, Eunae Kim², Abigail A. Wise¹, Patrick J. O'Brien³, Ioan Andricioaei² & Hashim M. Al-Hashimi¹

Sequence-directed variations in the canonical DNA double helix structure that retain Watson–Crick base-pairing have important roles in DNA recognition, topology and nucleosome positioning. By using nuclear magnetic resonance relaxation dispersion spectroscopy in concert with steered molecular dynamics simulations, we have observed transient sequence-specific excursions away from Watson–Crick base-pairing at CA and TA steps inside canonical duplex DNA towards low-populated and short-lived A•T and G•C Hoogsteen base pairs. The observation of Hoogsteen base pairs in DNA duplexes specifically bound to transcription factors and in damaged DNA sites implies that the DNA double helix intrinsically codes for excited state Hoogsteen base pairs as a means of expanding its structural complexity beyond that which can be achieved based on Watson–Crick base-pairing. The methods presented here provide a new route for characterizing transient low-populated nucleic acid structures, which we predict will be abundant in the genome and constitute a second transient layer of the genetic code.

Soon after its discovery¹, it was recognized that the DNA double helix could accommodate a range of conformations that retain Watson–Crick (WC) base-pairing². Sequence-directed variations in duplex DNA structure, shape and flexibility have since been shown to have fundamental roles in biology, including in the indirect readout of DNA sequences by recognition factors^{3,4}, nucleosome positioning^{4,5} and formation of loops and other large-scale architectures⁶ involved in DNA packaging, replication, transcription and recombination. DNA duplexes resiliently maintain WC base-pairing even when supercoiled and wrapped around histone octamers in nucleosomes⁷ and when adopting left-handed double-helical conformations known as Z-DNA⁸. Deviations from the WC base-pairing have so far only been observed in duplex DNA bound to proteins^{9,10} and small molecule ligands^{11,12} and in the context of damaged DNA^{13,14}, but never within naked canonical B-DNA duplexes.

Thus far, atomic resolution structural studies of the iconic DNA double helix have exclusively focused on its dominant, experimentally accessible, ground state conformation. Far less is known about other low-energy DNA conformations that may be sampled only transiently in solution. NMR relaxation dispersion experiments^{15,16} provide a rare opportunity to detect and characterize such short-lived (< 5 ms) and low-populated species (> 0.1%), often referred to as 'excited states'. This methodology has been widely used to characterize protein excited states that have been implicated in folding^{15,16}, recognition¹⁷ and catalysis¹⁸, culminating in the recent structure determination of a transient protein-folding intermediate¹⁹. Recent advances in carbon-based relaxation dispersion experiments combined with selective labelling schemes have addressed limitations that have hindered application of this methodology to nucleic acids, allowing detection of excited states in both RNA^{20,21} and DNA^{21,22}. However, the structures of nucleic acid excited states remain elusive.

Excited states in CA steps of duplex DNA

We used a recently introduced carbon $R_{1\rho}$ relaxation dispersion NMR experiment that allows detection of excited states with enhanced time-scale sensitivity²¹ to probe for the existence of excited states in canonical

DNA duplexes (Fig. 1a). We uncovered chemical exchange processes directed towards excited states occurring specifically at A•T and G•C base pairs in CA/TG steps (Fig. 1b), which together with TA steps are the most flexible dinucleotide steps in DNA and frequently the confluence point for local structural deformations²³. In particular, we observed significant $R_{1\rho}$ relaxation dispersion, indicative of chemical exchange, at the base C8 and sugar C1' carbons of the adenine and guanine residues, and for the C6 carbon of the cytosine residue. No relaxation dispersion was observed for the adenine base C2, the thymine residue, or the cytosine sugar C1' sites (Fig. 1b, Supplementary Fig. 3). The exchange process is enhanced by longer 3' neighbouring A-tracts, modulated by positional context, and for the G•C base pair, is highly pH-dependent (Supplementary Fig. 2). A two-state analysis ($A \xrightleftharpoons[k_B]{k_A} B$)

of the off-resonance relaxation dispersion data (Fig. 1c, Supplementary Fig. 2) revealed a single base pair exchange process that is slightly faster for A•T versus G•C and that is directed towards minutely populated ($p_B \sim 0.64\%$ and $\sim 0.47\%$ for G•C and A•T in A₆-DNA, 26 °C) excited states that have exceptionally short lifetimes ($\tau_B = 1/k_B \sim 1.5$ ms and ~ 0.3 ms for G•C and A•T in A₆-DNA, 26 °C), and that have downfield-shifted carbon chemical shifts ($\Delta\omega_{AB(C8)} \sim 2.7$ – 3.2 p.p.m., $\Delta\omega_{AB(C1')} \sim 3.1$ – 3.7 p.p.m., and $\Delta\omega_{AB(C6)} \sim 2.2$ p.p.m., Supplementary Table 2).

Assignment to Hoogsteen base pairs

What is the excited state encoded by CA/TG steps and detected by NMR relaxation dispersion? NMR imino proton exchange measurements^{24,25} and computer simulations²⁶ have shown previously that WC base pairs can spontaneously break open and access extrahelical conformations. Whereas the forward exchange rates (k_A) measured by relaxation dispersion (~ 4 – 20 s⁻¹ at 26 °C, Supplementary Table 2) are within an order of magnitude of rates reported previously for base-pair opening (~ 40 – 400 s⁻¹ at 25 °C)²⁷, the excited states detected here are at least three orders of magnitude more populated (Supplementary Table 2) and point to more energetically favourable species. To obtain more insights into the excited state, we measured carbon $R_{1\rho}$

¹Department of Chemistry and Biophysics, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109, USA. ²Department of Chemistry, University of California, Natural Sciences 2, Irvine, California 92697, USA. ³Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA.

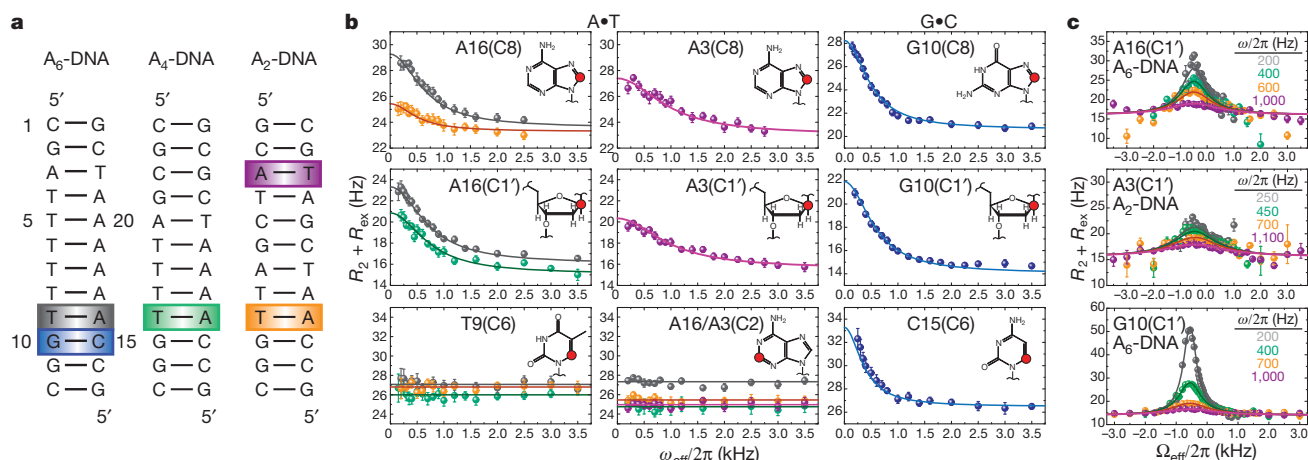


Figure 1 | Detection of base-pair-specific excited states in CA/TG steps of duplex DNA. **a**, DNA constructs containing varying length A-tracts with colour-coded A•T and G•C base pairs at CA/TG steps that show carbon chemical exchange. **b**, On-resonance ^{13}C $R_{1\rho}$ relaxation dispersion profiles for A•T (26.0 °C) and G•C (30.5 °C) showing CA/TG specific chemical exchange at purine base C8 and sugar C1' and at cytosine base C6. Shown are the best

base pair global fits (solid line) to a two-state asymmetric exchange model (Methods, equation 1). **c**, Representative off-resonance relaxation dispersion profiles for corresponding C1' sites and best global fits as in **b**. All error bars represent experimental uncertainty (one s.d.) estimated from mono-exponential fitting of duplicate sets of $R_{1\rho}$ data and analysis of signal-to-noise.

relaxation dispersion as a function of temperature (Fig. 2a). We then used transition-state theory and van't Hoff analysis to extract a complete thermodynamic-kinetic description of the two-state equilibria (Supplementary Table 5). Semi-logarithmic van't Hoff plots revealed a linear dependence characteristic of a two-state process for A•T and G•C base pairs (Fig. 2b). The analysis yielded activation free energies (~ 16 kcal mol $^{-1}$) and enthalpies (~ 12 – 26 kcal mol $^{-1}$) for the forward transition (Fig. 2c) that are comparable to values reported previously for base-pair opening (~ 14 – 20 kcal mol $^{-1}$ and ~ 8 – 29 kcal mol $^{-1}$ respectively) 27,28 . Thus, the transition to the excited state entails disruption of stacking and hydrogen-bonding interactions in the WC base pair. However, for both A•T and G•C, this loss in enthalpy is nearly entirely restored when the excited state is formed. In fact, the excited state is in part destabilized relative to the WC ground state by a less favourable entropy, implying a more 'rigid' excited state conformation and counter to what would be expected for a flexible looped-out state.

Taken together, our data point to an excited state conformation whose creation requires complete disruption of WC base pairs, but

whose thermodynamic stability is comparable to that of a WC base pair. One possibility is that the excited state represents an alternative base pair. Here, the correlated nature of the exchange at purine C8 and C1' nuclei and the large downfield carbon chemical shifts provides important clues. In particular, the magnitude and direction of $\Delta\omega_{\text{AB(C8)}}$ and $\Delta\omega_{\text{AB(C1')}}$ are strongly indicative of an *anti-to-syn* transition, as deduced from a survey of carbon chemical shifts 29 and density functional theory (DFT) calculations 30 . Remarkably, an *anti-to-syn* rotation of the adenine or guanine base results in creation of a Hoogsteen (HG) base pair optimally stabilized by two hydrogen bonds (Fig. 3a). The HG G•C base pair would require protonation of cytosine N3 (G•C $^+$, pK $_a$ (N3) ~ 4.2 , ref. 31), which can in turn explain the relaxation dispersion observed at cytosine C6 at neutral and acidic pH, and the downfield shifted excited state C6 chemical shift, as well as the pH dependence of dispersion measured at carbon sites in the G•C but not in the A•T base pair (see Supplementary Discussion). HG base pairs have been widely observed in non-canonical DNA structures, such as DNA triplexes and quadruplexes, where they

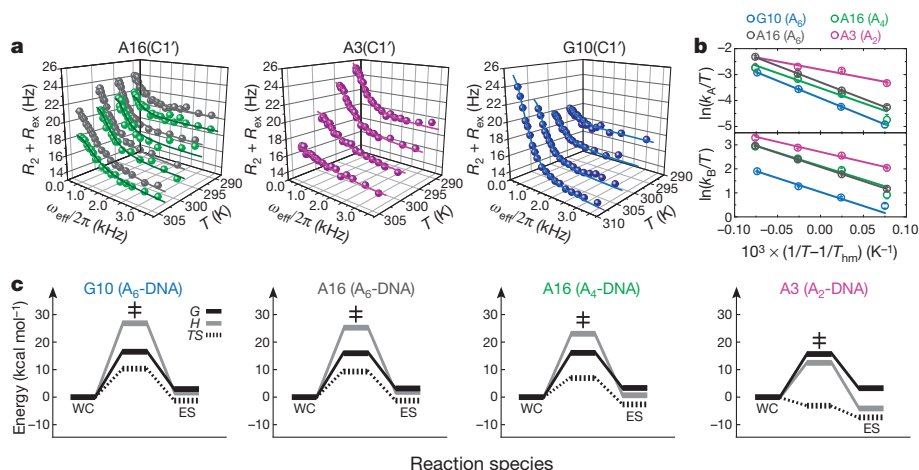


Figure 2 | Kinetic-thermodynamic analysis of ground-to-excited state transitions. **a**, Representative on-resonance ^{13}C $R_{1\rho}$ relaxation dispersion profiles as a function of temperature for A16 (A₆-DNA and A₄-DNA), A3 (A₂-DNA), and G10 (A₆-DNA) C1'. **b**, Modified van't Hoff plots showing temperature dependence of the forward (k_{A}) and reverse (k_{B}) rate constants for the two-site exchange in A•T and G•C base pairs highlighted in Fig. 1a. Error

bars represent experimental uncertainty (one s.d.) as determined from propagation of errors obtained from mono-exponential fitting of duplicate sets of $R_{1\rho}$ data and analysis of signal-to-noise. **c**, Corresponding kinetic-thermodynamic profiles for exchange between the WC ground state and the excited state (ES) via a transition state (\ddagger), showing activation and net free energy (G), enthalpy (H), and entropy (TS) changes (referenced to 0).

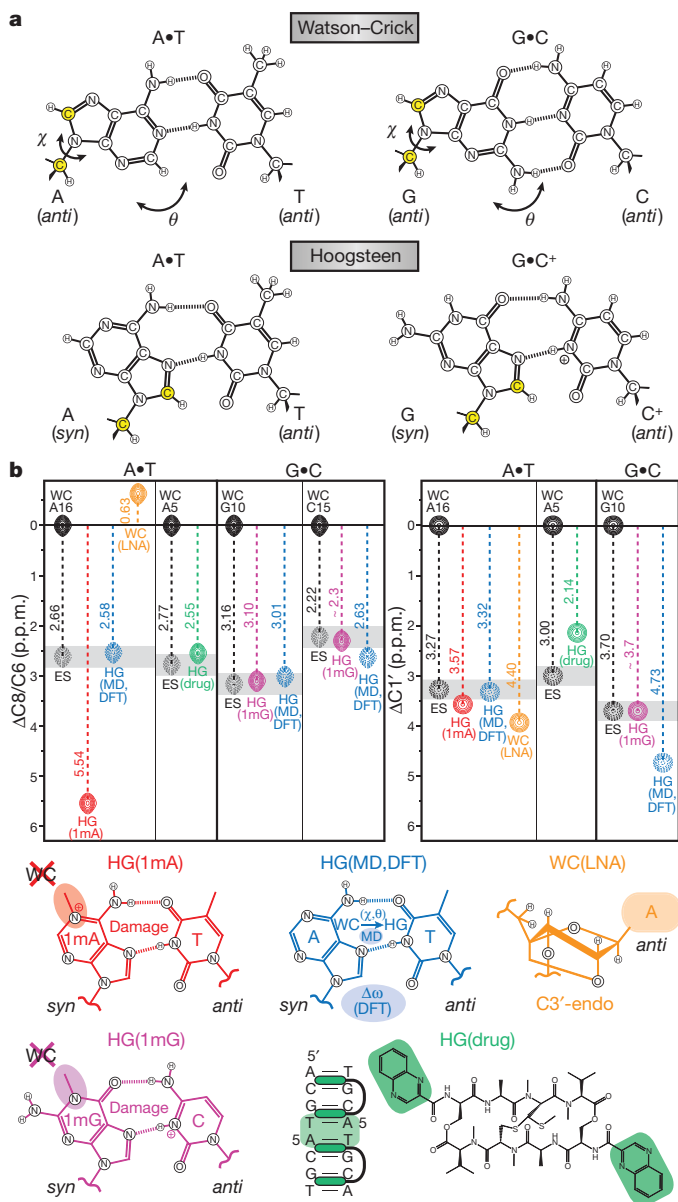


Figure 3 | Chemical shift assignment of excited state Hoogsteen base pairs.
a, Chemical structures for WC and HG A•T and G•C⁺ base pairs. The HG geometry can be achieved by purine rotation around the glycosidic bond (χ) and base-flipping (θ), affecting simultaneously C8 and C1' (yellow).
b, Depiction of C8 and C1' chemical shifts relative to WC for the excited state (ES, grey); an N1-methyladenine modified A₆-DNA (HG(1mA), red), an N1-methylguanine modified A₆-DNA (HG(1mG), violet), and an echinomycin-bound DNA (HG(drug), green) forming HG base pairs; a simulated A16•T9 or G10•C15 HG base pair in A₆-DNA (HG(MD, DFT), blue); a C3'-endo locked A₆-DNA (WC(LNA), orange); and representative cartoons.

are specifically recognized by proteins³². In a few cases, HG base pairs have also been observed in duplex DNA containing alternating AT repeats³³ or in complex with ligands^{11,12} and proteins, including the active site of DNA polymerase- α ³⁴ and a complex between TATA binding protein (TBP) and a mutant TATA-box DNA⁹, where HG G•C⁺ base pairs are formed at near neutral pH. Indeed, HG base pairs in an all-*anti* DNA triplex show similar stability to that normally found for WC base pairs in duplex DNA³⁵. An excited-state HG base pair can explain the lower enthalpy and lower entropy of the excited state relative to a base pair open state. Its creation requires a ~ 180 -degree base rotation around the glycosidic linkage and disruption of the WC base pair, consistent with our measured transition state barriers. These

excited state HG base pairs may have evaded detection by prior solvent exchange measurements due to hydrogen-bond protection of the thymine imino proton in (A•T) and/or inaccessibility of the guanine imino proton in (G•C) in addition to slower exchange rates. For example, the observation of a non-hydrogen bonded G H1 resonance in a G(*syn*)•A⁺(*anti*) mismatch base pair at 20 °C inside a DNA duplex³⁶ supports the lower exchange rates with solvent.

Trapping HG base pairs in duplex DNA

To test further for the existence of HG base pairs, we used chemical modifications to trap the excited state HG base pair within duplex DNA. This allowed us to compare the carbon chemical shift signatures of the trapped HG base pair directly with those measured for the excited state using relaxation dispersion. By installing an *N*1-methylated adenine (1mA), which is a common DNA lesion known to impair WC base pairs sterically and favour HG base pairs^{13,37} at the CA step, we trapped the HG A•T base pair in the A₆-DNA duplex as confirmed by analysis of nuclear Overhauser effect (NOE) connectivity and proton chemical shifts (Supplementary Fig. 5). Though never observed previously, we also successfully trapped an HG G•C⁺ base pair using instead the structural analogue to 1mA—an *N*1-methylated guanine (1mG)—at the same CA step (Supplementary Fig. 5). The ability to trap HG A•T or G•C⁺ base pairs independently is consistent with the relaxation dispersion data showing that transitions to excited state A•T and G•C HG base pairs within CA steps are semi-independent of one another. Strikingly, the differences in carbon chemical shift measured between modified and unmodified DNA were in excellent agreement with the chemical shift differences measured between the ground and excited state by relaxation dispersion (Fig. 3b). The noticeably downfield chemical shift observed for the trapped 1mA versus the unmodified excited state HG base pair could be attributed to changes in the electronic environment arising from introduction of the methyl group and positive charge (not present in 1mG), as supported by DFT calculations (Supplementary Information) and possibly from slight differences in the HG geometry. Conversely, resonances that showed small differences in carbon chemical shift between modified and unmodified constructs (Supplementary Table 4) exhibited little to no carbon chemical exchange (Supplementary Fig. 3). Our data rule out other non-canonical base-pairing modes involving pyrimidine base rotation or a repuckered sugar conformation from C2'-endo to C3'-endo, which yields upfield rather than downfield nucleobase carbon chemical shifts (Fig. 3b and Supplementary Fig. 4). Thus, comparison of known HG ground state chemical shifts with the excited state chemical shifts provides strong support for its assignment as an HG base pair.

As an inverse experiment, we asked whether TA steps, which have also been observed to form HG base pairs in duplex DNA bound to small molecule ligands^{11,12}, exhibit the characteristic HG excited state. We measured carbon relaxation dispersion data for a palindromic DNA sequence, which has previously been shown to form tandem A•T HG base pairs in solution³⁸ when in complex with the bis-intercalating antibiotic echinomycin (Fig. 3b). Strikingly, we observe the same chemical exchange pattern in the TA step and carbon chemical shift differences between ground and excited state that are in excellent agreement with the differences between the free (WC) DNA and drug-bound (HG) DNA (Fig. 3b and Supplementary Fig. 6). Thus, excited state HG base pairs are not restricted to CA/TG steps but also include TA steps, and they can be conformationally captured by recognition factors. In fact, our recent observation of chemical exchange in a TA step containing an 1,N6-ethenoadenine adduct²¹ could potentially be explained by transient *anti*-to-*syn* excursions as in the HG base pair. Moreover, the weaker binding affinity of TBP to an HG-containing mutant TATA box, where a WC base pair would not be tolerated, versus the wild-type sequence with all-WC base-pairing ($\Delta\Delta G \sim 3 \text{ kcal mol}^{-1}$)³⁹ could be attributed to conformational selection of a low-populated HG base pair and could be further correlated with transcriptional regulation of gene repression⁴⁰.

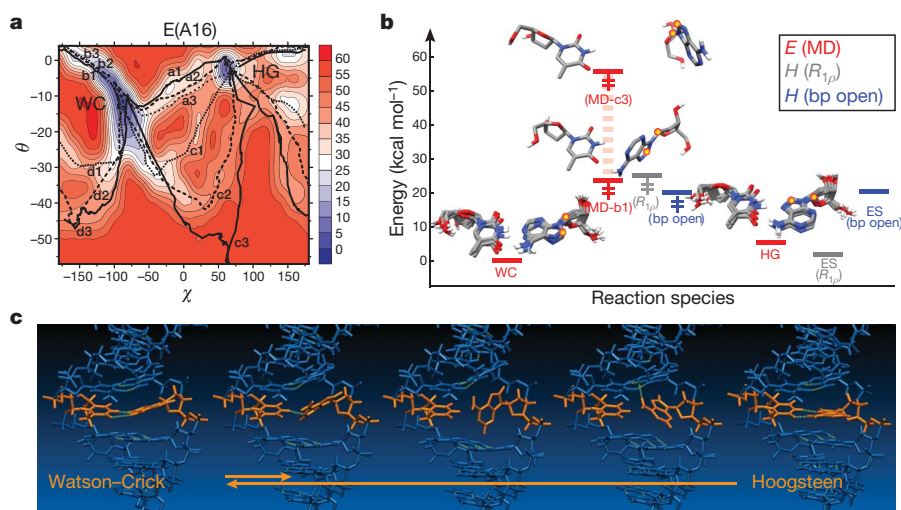


Figure 4 | Watson-Crick to Hoogsteen base pair transition simulations. **a**, Pseudo-free-energy (E , kcal mol⁻¹) contour plots as a function of (χ , θ) pairs for A16•T9 obtained from multiple CPR trajectories (a–d). **b**, Initial WC and final excited-state (ES) HG structures and representative lowest-energy (b1) and highest-energy (c3) transition state (\ddagger) structures of A16•T9 illustrating the

span of CPR transition barriers, and their relative potential energies (averaged for WC and HG) compared with enthalpies (H) derived from NMR $R_{1\rho}$ relaxation dispersion for chemical exchange or NMR imino proton exchange for A•T base pair opening²⁷. **c**, Snapshots from a representative CPR transition pathway (a1) for A16•T9.

Simulations of WC-to-HG transition

To assess the energetic and stereochemical feasibility of the proposed WC-to-HG transition, as well as to obtain insights into the transition pathway, we used conjugate peak refinement (CPR)⁴¹ methods to simulate multiple transition pathways between WC and excited state HG base pairs that sample various glycosidic (χ) and base opening (θ) angles (Fig. 4a). CPR trajectories for A16•T9 and G10•C15 in A₆-DNA showed smooth WC-to-HG transitions via *anti*-to-*syn* purine rotation around χ accompanied by minor adjustments in the complementary pyrimidine residue and neighbouring WC base pairs (Fig. 4a, Supplementary Fig. 7 and Supplementary Movies). Optimal pathways feature purine base-flips at low base pair opening angles into the major groove and transition states, which require complete disruption of WC base-pairing—this is also observed experimentally by relaxation dispersion—in which the purine base is near-orthogonal to its pyrimidine partner. This unusual geometry is well accommodated by the rather large inter-base-pair spacing in the B-form helix (that is, base pair rise ~ 3.3 Å) and its ability to mould without loss of other WC base-pairing. Although a large spread of transition barriers are sampled, the minimal energy barriers and net energy changes are within 2 kcal mol⁻¹ of NMR-derived ensemble enthalpy terms (Fig. 4b). The final HG state for G10•C15 yields generally lower relative CPR energies—it is observed to adopt variable *syn*-guanine geometry, with either one or no optimal hydrogen bonds to the cytosine due to the lack of N3 protonation that stabilizes the ideal HG base pair, but on occasion forming an intra-residue hydrogen bond to a backbone oxygen (Supplementary Movie 2).

The carbon chemical shifts computed for the WC and lowest-energy HG geometries using DFT also exhibited the characteristic downfield-shifted $\Delta\omega_{AB(C8)}$ and $\Delta\omega_{AB(C1')}$ observed experimentally by relaxation dispersion (Fig. 3b and Supplementary Table 4). Conversely, the predicted chemical shifts revealed little and/or random variations between WC and HG base pairs for adenine C2, thymine C6 and cytosine C1', consistent with the lack of observable relaxation dispersion at those sites (Supplementary Fig. 8). Comprehensive DFT calculation of carbon chemical shifts for conformers sampled in various pathways between WC and HG states (Supplementary Fig. 8) yielded a small number of geometries that match the measured excited-state chemical shifts, but that could readily be excluded because they involve barrierless transitions and/or represent high-energy structures that disagree with experimentally derived enthalpies.

Our data strongly indicate that the DNA double helix codes for a preexisting WC-to-HG equilibrium, with HG base pairs representing an accessible and energetically competent alternative to WC base-pairing that present very distinct electrostatic and hydrophobic signatures. This makes it possible to trap HG base pairs by interactions with cellular triggers, thereby expanding the structural and functional diversity of the double helix beyond that which can be achieved based on an alphabet of only WC base-pairing. There are several examples of transcription factors including TBP⁹ and p53 tumour suppressor¹⁰ that specifically recognize HG base pairs embedded in different WC contexts, where the modulation in binding affinity, conceivably, by an HG base pair could even be correlated with an essential biological function⁴⁰. In addition, HG base pairs are often trapped by oxidative and alkylation lesions¹⁴, providing unique recognition signals for repair enzymes in search of damage sites in a sea of undamaged DNA. Transient formation of HG base pairs inside B-DNA may also serve to promote non-canonical structures such as contiguous HG motifs, especially in tandem CA and TA repeats, or more dramatic transformations to Z-DNA⁴², and may well exist in much greater abundance for native genomic DNA, which is under torsional stress in the cellular environment. The methods presented here provide a general strategy for detecting and characterizing excited states in DNA and RNA, which we predict will be abundant in the genome and constitute another transient layer of the genetic code.

METHODS SUMMARY

Detailed methods on DNA sample preparation and assignment, NMR relaxation dispersion data collection and analysis, MD/CPR simulations, and DFT chemical shift calculations can be found in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 December 2010; accepted 5 January 2011.

Published online 26 January 2011.

- Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
- Record, M. T. Jr *et al.* Double helical DNA: conformations, physical properties, and interactions with ligands. *Annu. Rev. Biochem.* **50**, 997–1024 (1981).
- Koudelka, G. B., Mauro, S. A. & Ciubotaru, M. Indirect readout of DNA sequence by proteins: the roles of DNA sequence-dependent intrinsic and extrinsic forces. *Prog. Nucleic Acid Res. Mol. Biol.* **81**, 143–177 (2006).
- Rohs, R. *et al.* The role of DNA shape in protein–DNA recognition. *Nature* **461**, 1248–1253 (2009).

5. Segal, E. *et al.* A genomic code for nucleosome positioning. *Nature* **442**, 772–778 (2006).
6. Saiz, L. & Vilar, J. M. DNA looping: the consequences and its control. *Curr. Opin. Struct. Biol.* **16**, 344–350 (2006).
7. Richmond, T. J. & Davey, C. A. The structure of DNA in the nucleosome core. *Nature* **423**, 145–150 (2003).
8. Wang, A. H. *et al.* Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **282**, 680–686 (1979).
9. Patikoglou, G. A. *et al.* TATA element recognition by the TATA box-binding protein has been conserved throughout evolution. *Genes Dev.* **13**, 3217–3230 (1999).
10. Kitayner, M. *et al.* Diversity in DNA recognition by p53 revealed by crystal structures with Hoogsteen base pairs. *Nature Struct. Mol. Biol.* **17**, 423–429 (2010).
11. Ughetto, G. *et al.* A comparison of the structure of echinomycin and triostin A complexed to a DNA fragment. *Nucleic Acids Res.* **13**, 2305–2323 (1985).
12. Seaman, F. C. & Hurley, L. Interstrand cross-linking by bizelesin produces a Watson-Crick to Hoogsteen base-pairing transition region in d(CGTAATTACG)₂. *Biochemistry* **32**, 12577–12585 (1993).
13. Yang, H., Zhan, Y., Fenn, D., Chi, L. M. & Lam, S. L. Effect of 1-methyladenine on double-helical DNA structures. *FEBS Lett.* **582**, 1629–1633 (2008).
14. Shanmugam, G., Kozekov, I. D., Guengerich, F. P., Rizzo, C. J. & Stone, M. P. Structure of the 1,N²-ethenodeoxyguanosine adduct opposite cytosine in duplex DNA: Hoogsteen base pairing at pH 5.2. *Chem. Res. Toxicol.* **21**, 1795–1805 (2008).
15. Palmer, A. G. III. NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.* **104**, 3623–3640 (2004).
16. Korzhnev, D. M. & Kay, L. E. Probing invisible, low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding. *Acc. Chem. Res.* **41**, 442–451 (2008).
17. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chem. Biol.* **5**, 789–796 (2009).
18. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
19. Korzhnev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R. & Kay, L. E. A transient and low-populated protein-folding intermediate at atomic resolution. *Science* **329**, 1312–1316 (2010).
20. Johnson, J. E. Jr & Hoogstraten, C. G. Extensive backbone dynamics in the GCAA RNA tetraloop analyzed using ¹³C NMR spin relaxation and specific isotope labeling. *J. Am. Chem. Soc.* **130**, 16757–16769 (2008).
21. Hansen, A. L., Nikolova, E. N., Casiano-Negroni, A. & Al-Hashimi, H. M. Extending the range of microsecond-to-millisecond chemical exchange detected in labeled and unlabeled nucleic acids by selective carbon R_{1ρ} NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 3818–3819 (2009).
22. Shajani, Z. & Varani, G. ¹³C relaxation studies of the DNA target sequence for HhaI methyltransferase reveal unique motional properties. *Biochemistry* **47**, 7617–7625 (2008).
23. Travers, A. A. The structural basis of DNA flexibility. *Philos. Transact. A Math. Phys. Eng. Sci.* **362**, 1423–1438 (2004).
24. Pardi, A., Morden, K. M., Patel, D. J. & Tinoco, I. Jr. Kinetics for exchange of imino protons in the d(C-G-C-G-A-A-T-T-C-G-C-G) double helix and in two similar helices that contain a G•T base pair, d(C-G-T-G-A-A-T-T-C-G-C-G), and an extra adenine, d(C-G-C-A-G-A-A-T-T-C-G-C-G). *Biochemistry* **21**, 6567–6574 (1982).
25. Guéron, M., Kochoyan, M. & Leroy, J. L. A single mode of DNA base-pair opening drives imino proton exchange. *Nature* **328**, 89–92 (1987).
26. Pérez, A., Luque, F. J. & Orozco, M. Dynamics of B-DNA on the microsecond time scale. *J. Am. Chem. Soc.* **129**, 14739–14745 (2007).
27. Coman, D. & Russu, I. M. A nuclear magnetic resonance investigation of the energetics of basepair opening pathways in DNA. *Biophys. J.* **89**, 3285–3292 (2005).
28. Song, K. *et al.* An improved reaction coordinate for nucleic acid base flipping studies. *J. Chem. Theory Comput.* **5**, 3105–3113 (2009).
29. Greene, K. L., Wang, Y. & Live, D. Influence of the glycosidic torsion angle on ¹³C and ¹⁵N shifts in guanosine nucleotides: investigations of G-tetrad models with alternating *syn* and *anti* bases. *J. Biomol. NMR* **5**, 333–338 (1995).
30. Xu, X. & Au-Yeung, S. Investigation of chemical shift and structure relationships in nucleic acids using NMR and density functional theory methods. *J. Phys. Chem. B* **104**, 5641–5650 (2000).
31. Izatt, R. M., Christensen, J. J. & Rytting, J. H. Sites and thermodynamic quantities associated with proton and metal ion interaction with ribonucleic acid, deoxyribonucleic acid, and their constituent bases, nucleosides, and nucleotides. *Chem. Rev.* **71**, 439–481 (1971).
32. Ghosal, G. & Muniyappa, K. Hoogsteen base-pairing revisited: resolving a role in normal biological processes and human diseases. *Biochem. Biophys. Res. Commun.* **343**, 1–7 (2006).
33. Abrescia, N. G., Thompson, A., Huynh-Dinh, T. & Subirana, J. A. Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing. *Proc. Natl Acad. Sci. USA* **99**, 2806–2811 (2002).
34. Nair, D. T., Johnson, R. E., Prakash, L., Prakash, S. & Aggarwal, A. K. Human DNA polymerase ϵ incorporates dCTP opposite template G via a G.C+ Hoogsteen base pair. *Structure* **13**, 1569–1577 (2005).
35. Powell, S. W., Jiang, L. & Russu, I. M. Proton exchange and base pair opening in a DNA triple helix. *Biochemistry* **40**, 11065–11072 (2001).
36. Sau, A. K. *et al.* Evidence for A⁺ (*anti*)-G(*syn*) mismatched base-pairing in d-GGTAAGCGTACC. *FEBS Lett.* **377**, 301–305 (1995).
37. Lu, L., Yi, C., Jian, X., Zheng, G. & He, C. Structure determination of DNA methylation lesions N¹-meA and N²-meC in duplex DNA using a cross-linked protein-DNA system. *Nucleic Acids Res.* **38**, 4415–4425 (2010).
38. Gilbert, D. E., van der Marel, G. A., van Boom, J. H. & Feigon, J. Unstable Hoogsteen base pairs adjacent to echinomycin binding sites within a DNA duplex. *Proc. Natl Acad. Sci. USA* **86**, 3006–3010 (1989).
39. Hoopes, B. C., LeBlanc, J. F. & Hawley, D. K. Contributions of the TATA box sequence to rate-limiting steps in transcription initiation by RNA polymerase II. *J. Mol. Biol.* **277**, 1015–1031 (1998).
40. Meyer, T., Carlstedt-Duke, J. & Starr, D. B. A weak TATA box is a prerequisite for glucocorticoid-dependent repression of the osteocalcin gene. *J. Biol. Chem.* **272**, 30709–30714 (1997).
41. Fischer, S. & Karplus, M. Conjugate peak refinement: an algorithm for finding reaction paths and accurate transition-states in systems with many degrees of freedom. *Chem. Phys. Lett.* **194**, 252–261 (1992).
42. Segers-Nolten, G. M., Sijtsma, N. M. & Otto, C. Evidence for Hoogsteen GC base pairs in the proton-induced transition from right-handed to left-handed poly(dG-dC)·poly(dG-dC). *Biochemistry* **36**, 13241–13247 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. L. Hansen, S. Horowitz and J. Feigon for valuable discussions and suggestions, A. V. Kurochkin for NMR expertise, and C.L. Brooks III for access to a supercomputing cluster. We gratefully acknowledge the Michigan Economic Development Cooperation and the Michigan Technology Tri-Corridor for support in the purchase of a 600 MHz spectrometer. This work was supported by NSF CAREER awards (MCB 0644278 to H.M.A. and CHE-0918817 to I.A.) and an NIH grant (R01GM089846). E.N.N. acknowledges support by a Rackham International and Predoctoral Fellowship awarded by the University of Michigan.

Author Contributions E.N.N. prepared DNA samples assisted by A.A.W. and performed/analysed all NMR experiments and DFT calculations; E.N.N. and H.M.A. conceived the idea of an HG excited state base pair and approaches to investigate its formation; I.A. and E.K. performed and analysed the MD/CPR simulations; P.J.O. provided expertise and guidance for damaged DNA studies along with critical manuscript revisions; H.M.A. and E.N.N. with help from P.J.O., E.K. and I.A. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to H.M.A. (hashimi@umich.edu) or I.A. (andricio@uci.edu).

METHODS

Preparation and NMR resonance assignment of unlabelled and $^{13}\text{C}/^{15}\text{N}$ -labelled DNA. Isotopically labelled DNA dodecamers (Fig. 1a) were synthesized by *in vitro* primer extension using a template hairpin DNA (IDT), Klenow fragment DNA polymerase (NEB), and uniformly $^{13}\text{C}/^{15}\text{N}$ -labelled dNTPs (Isotec, Sigma-Aldrich). Single-stranded DNA products were purified by 20% denaturing polyacrylamide gel electrophoresis (PAGE), isolated by passive elution and desalted on a C18 reverse-phase column (Sep-pak, Waters). Oligonucleotides were lyophilized and resuspended in NMR buffer (15 mM sodium phosphate pH 6.8, 25 mM sodium chloride, 0.1 mM EDTA, 10% D_2O). Complementary oligonucleotides were annealed at an equimolar ratio typically at 0.5–1.0 mM for NMR studies. Unlabelled DNA oligonucleotides were purchased from IDT (A_2 -DNA, A_4 -DNA, A_6 -DNA in Fig. 1a, and E-DNA in Supplementary Fig. 6), Exiqon, Inc. (A_6 -DNA^{A16LNA} and A_2 -DNA^{A16LNA} in Supplementary Fig. 4) and Midland Certified, Inc. (A_6 -DNA^{A16} and A_6 -DNA^{A16G10} in Supplementary Fig. 5). Unlabelled DNA constructs including equivalent samples to $^{13}\text{C}/^{15}\text{N}$ -labelled DNA were prepared as described²¹ at 2.0–4.0 mM concentrations and assigned using conventional ^1H - ^1H nuclear Overhauser effect spectroscopy (NOESY) in 10% D_2O at 5 °C and/or 26 °C. The 2:1 complex between echinomycin (Selleck Chemicals) and E-DNA was prepared as described previously³⁸. All NMR experiments were performed on a Bruker Avance 600 MHz NMR spectrometer equipped with a 5 mm triple-resonance cryogenic probe.

Selective ^{13}C $R_{1\rho}$ relaxation dispersion. Rotating frame ($R_{1\rho}$) relaxation dispersion profiles were measured at a single field (14.1 T) using a selective carbon experiment with a one-dimensional acquisition scheme²¹ that extends the sensitivity to chemical exchange into millisecond timescales relative to conventional two-dimensional relaxation dispersion methods. On-resonance data were recorded at variable (100 to 3,500 Hz) effective spinlock field strength (ω_{eff}) (Supplementary Tables 1 and 3) for various sites in $^{13}\text{C}/^{15}\text{N}$ -labelled and unlabelled DNA constructs. For $^{13}\text{C}/^{15}\text{N}$ -labelled DNA samples (Supplementary Fig. 2) and E-DNA octamer (Supplementary Fig. 6), off-resonance dispersion data were collected in a temperature dependent manner at variable spinlock offset frequencies ($\Omega_{\text{eff}} \approx \Omega_{\text{ave}}$, see below) and at three to four different spinlock powers (ω_1) (Supplementary Table 1 and 3). In each case, the following relaxation delays were used: {0, 4, 8, 12 (2×), 16, 20, 26, 32 (2×) ms} for C2/C6/C8 and {0, 4, 8, 12 (2×), 18, 26, 34, 42 (2×) ms} for C1' in $^{13}\text{C}/^{15}\text{N}$ -labelled DNA constructs; {0, 40 (2×) ms} for C8 and {0, 48 (2×) ms} for C1' at 17 °C; {0, 48 (2×) ms} for C8 and {0, 60 (2×) ms} for C1' at 26 °C in E-DNA; {0, 30 (2×) ms} for C8 in A_6 -DNA^{A16LNA}. Data points corresponding to Hartmann–Hahn matching conditions were omitted from the data fits as described previously²¹. Data were processed using nmPipe⁴³ and the $R_{1\rho}$ or the effective transverse relaxation rates ($R_{2\text{eff}} = R_2 + R_{\text{ex}}$) were computed by fitting the resonance intensities to monoexponential decays using Mathematica 6.0 (Wolfram Research).

Measured relaxation dispersion profiles with on- and off-resonance data were fit by equation (1) (below) that assumes a two-state equilibrium ($\text{A} \xrightleftharpoons[k_{\text{B}}]{k_{\text{A}}} \text{B}$) with an asymmetric population distribution ($p_{\text{A}} \gg p_{\text{B}}$)⁴⁴ using Mathematica 6.0 (Wolfram Research):

$$R_{1\rho} = R_1 \cos^2 \theta + R_2 \sin^2 \theta + \sin^2 \theta \frac{p_{\text{A}} p_{\text{B}} \Delta\omega_{\text{AB}}^2 k_{\text{ex}}}{\Omega_{\text{B}}^2 + \omega_1^2 + k_{\text{ex}}^2} \quad (1)$$

where R_1 and R_2 are the intrinsic longitudinal and transverse relaxation rates, respectively (assumed to be identical for A and B species), Ω is the resonance offset from the spinlock carrier, ω_1 is the spinlock strength; $\tan \theta = \omega_1 / \Omega_{\text{ave}}$, $\Delta\omega_{\text{AB}} = \Omega_{\text{B}} - \Omega_{\text{A}}$, $\Omega_{\text{ave}} = p_{\text{A}} \Omega_{\text{A}} + p_{\text{B}} \Omega_{\text{B}}$, where p_{A} (p_{B}) is the major (minor) state fractional population ($p_{\text{A}} + p_{\text{B}} = 1$); $k_{\text{ex}} = k_{\text{A}} + k_{\text{B}}$ is the exchange rate constant for a two-state equilibrium, where $k_{\text{A}} = p_{\text{B}} k_{\text{ex}}$ and $k_{\text{B}} = p_{\text{A}} k_{\text{ex}}$ are the forward and reverse rate constants respectively. Similar results were obtained when fitting relaxation dispersion profiles against more complex two-state exchange models including the Laguerre approximation⁴⁵ that do not assume a skewed population distribution (data not shown) and statistical analysis implied that application of the more complex model here was not justified. Temperature-dependent data for base/sugar resonances within the same nucleotide or base pair were fit individually and globally with shared parameters (k_{ex} and p_{B} for each temperature) (Supplementary Tables 2 and 3). The best-fit model was assessed using F-test statistics (data not shown), which uses chi-square (χ^2) and F-distribution analysis to determine the feasibility of a complicated model (that is, individual fits) versus a more simple model (that is, shared-parameter fits) nested inside the first model. The chemical shift difference $\Delta\omega_{\text{AB}}$ was assumed to be invariant over the narrow temperature range investigated. On-resonance $R_{1\rho}$ profiles were fit by a simplified two-state fast exchange expression⁴⁴ ($k_{\text{ex}} \gg \Delta\omega_{\text{AB}}$):

$$R_{1\rho} = R_2 + R_{\text{ex}} = R_2 + \frac{\Phi_{\text{ex}} k_{\text{ex}}}{\omega_1^2 + k_{\text{ex}}^2}; \Phi_{\text{ex}} = p_{\text{A}} p_{\text{B}} \Delta\omega_{\text{AB}}^2 \quad (2)$$

where all parameters are as described above.

Thermodynamic analysis. The observed temperature dependence of the forward (k_{A}) and reverse (k_{B}) rate constants (Supplementary Table 2) was fit by a modified van't Hoff equation that accounts for statistical compensation effects and assumes a smooth energy surface²⁷:

$$\ln \left(\frac{k_i(T)}{T} \right) = \ln \left(\frac{k_{\text{B}} \kappa}{h} \right) - \frac{\Delta G_i^{\ddagger}(T_{\text{hm}})}{RT_{\text{hm}}} - \frac{\Delta H_i^{\ddagger}}{R} \left(\frac{1}{T} - \frac{1}{T_{\text{hm}}} \right) \quad (3)$$

where k_i ($i = \text{A}, \text{B}$) is the rate constant, ΔG_i^{\ddagger} and ΔH_i^{\ddagger} are the free energy and enthalpy of activation respectively, R is the universal gas constant, T is temperature, and T_{hm} is the harmonic mean of the experimental temperatures computed as $T_{\text{hm}} = n / \sum_{i=1}^n (1/T_i)$; k_{B} is Boltzmann's constant, h is Planck's constant, κ is the transmission coefficient (assumed to be 1) in the pre-exponential factor of Eyring's theory. The entropy of activation (ΔS_i^{\ddagger}) was calculated as follows:

$$T \Delta S_i^{\ddagger} = \Delta H_i^{\ddagger} - \Delta G_i^{\ddagger}(T_{\text{hm}}) \quad (4)$$

Semi-logarithmic plots are included in Fig. 2b and best-fit thermodynamic parameters are reported in Supplementary Table 5.

An alternative interpretation of the thermodynamic parameters is given by the phenomenological Ferry law^{46,47}, which incorporates a lower energy barrier with a rough enthalpic surface:

$$\ln(k_i) = \ln C - \frac{\Delta H_i^{\ddagger}}{RT} - \frac{\langle H_i^{\ddagger} \rangle}{(RT)^2} \quad (5)$$

where C is a constant, where $\langle H_i^{\ddagger} \rangle^{1/2}$ represents the enthalpy due to ruggedness of the barrier. The maximum $\langle H_i^{\ddagger} \rangle^{1/2}$ values were calculated by taking the smooth Arrhenius-like enthalpic barrier to be vanishing ($\Delta H_i^{\ddagger} = 0$) and are reported in Supplementary Table 4.

Molecular dynamics simulations of WC-to-HG base pair transition pathways.

An initial duplex DNA in standard B-form with Watson–Crick (WC) base-pairing corresponding to A_6 -DNA was generated using 3DNA⁴⁸. Hoogsteen (HG) base pairs were generated at A16•T9, A3•T22 and G10•C15 positions using known X-ray coordinates^{49,50}, where the purine adopts a *syn* conformation, whereas the complementary pyrimidine retains an *anti* conformation (see Fig. 3a). Initially, all conformers with WC or HG base-pairing were equilibrated through a series of energy minimizations using the Adopted Basis Newton–Raphson (ABNR) algorithm. Canonical (NVT) ensemble MD were performed with the CHARMM27 all-atom force field⁵¹ and the generalized Born molecular volume (GBMV) implicit solvation model (GBMV II)^{52–54}. The velocity–Verlet algorithm was used with a time step of 2 fs. A temperature of 300 K was kept constant with a Nosé–Hoover thermostat^{55,56}. The cutoff for non-bonded list generation was set to 21 Å, the cutoff for non-bonded interactions was set to 18 Å, and the onset of switching for non-bonded interactions was set to 16 Å. The SHAKE algorithm was used to constrain vibrations of covalent bonds to hydrogen atoms involved. To conserve the duplex DNA structure during pre-equilibration, flat-bottom distance restraints were applied, which prevented the hydrogen-bond donor from moving more than 2.0 Å. The simulations ran for a total of 6.0 ns (Supplementary Fig. 7).

For the collection and analysis of equilibrium data, the initial 1 ns simulation data were discarded. We constructed a two-dimensional free-energy map using the following two reaction coordinates: the DNA backbone root mean square deviation (r.m.s.d.) and the potential energy value; the minimization of this approximate free energy surface permitted us to choose the most populated structure in the canonical ensemble ($F = -k_{\text{B}} T \ln P(X, Y)$, X and Y are reaction coordinates, k_{B} is Boltzmann's constant, T is the absolute temperature (Supplementary Fig. 7)). The DNA backbone r.m.s.d. is defined by selecting the P-O5'-C5'-C4'-C3'-O3' atoms. The reference structure was a standard B-form DNA (twist angle $\Omega = 36.0^\circ$, rise per base pair along helix axis = 3.3 Å) without minimization. In order to probe the transition pathway between a WC and a HG base pair conformer, we used the conjugate peak refinement (CPR) method⁴¹ applicable to the study of complex isomerization reactions, including allosteric transitions in proteins and more general conformational changes in macromolecules. The resulting paths follow the adiabatic energy surface without applying any constraints and path-points between saddle-points ensure the continuity of the path, not necessarily constrained to find the absolute bottom of the energy valley. The initial WC-to-HG pathways were generated using a targeted molecular dynamics method that applied a holonomic constraint, which decreased gradually the r.m.s.d. to the final target structure. In each CPR cycle, a heuristic procedure was used to modify the path by improving, removing, or inserting one path-point, so that the new path avoids the maximum energy peak. Finally, to refine the CPR path further, we used a synchronous chain minimization method of all path-points, under the constraint that the points move within hyperplanes orthogonal to the path. The most populated structures with the WC base

pair or with the HG base pair from the normal MD simulations corresponded to 'reactant' and 'product' wells, respectively, on the energy surface. In order to sample a wider range of possible transition pathways, additional putative intermediates that differed in either the flip-over or flip-out angle were added by modifying, respectively, the glycosidic angle, χ (O4'-C1'-N9-C4), of the purine base and a centre-of-mass pseudo-dihedral angle, θ^{28} , which describes the extent to base opening. CPR data are reported in Supplementary Fig. 7.

Density functional theory calculations of carbon chemical shifts. DFT chemical shift calculations were conducted on a high-performance computing cluster using Gaussian 03 (ref. 57). DNA conformations of the target A•T and G•C base pair that represent a range of (χ, θ) pairs were selected from each simulated WC-to-HG transition pathway and capped by 3'OH/5'OH (UCSF Chimera⁵⁸) for DFT calculations (Supplementary Fig. 8) without further geometry optimization. NMR ^{13}C chemical shift calculations were conducted using the GIAO method with the B3LYP/6-311+G(2d,p) basis set. The isotropic carbon chemical shifts (σ_{ISO}) were referenced to TMS ($\sigma_{\text{TMS}} = 182.759$ p.p.m.) using the relationship $\delta_{\text{ISO}} = \sigma_{\text{TMS}} - \sigma_{\text{ISO}}$, where the structure of TMS was optimized at the same level of theory. Computed carbon chemical shifts were referenced to the most stable WC conformer in a given transition pathway and matched with NMR excited-state chemical shift differences ($\Delta\omega_{\text{AB}}$) in Supplementary Table 4. For benchmarking, similar DFT calculations were performed on single guanosine nucleotides with *anti* and *syn* glycosidic conformations from crystal structures of *Oxytricha nova* telomeric G-quadruplex (5'-(G)₄(T)₄(G)₄; PDB ID: IJPQ, IJRN, 2GWQ, 2GWE and 2NPR) with added hydrogen atoms (UCSF Chimera⁵⁸) and no further geometry optimization, and compared to observed NMR C8 chemical shifts for the same G-quadruplex (courtesy of M. Gill and P. Loria) (Supplementary Fig. 8). Similar calculations were performed for unmodified and N1-methyladenine modified HG base pairs and control WC base pairs using existing crystal structures (PDB ID: 1K61, 3IGK, 3IGL, 3KZ8, 3H8O and 2ATA).

43. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

44. Palmer, A. G. III & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev.* **106**, 1700–1719 (2006).
45. Miloushev, V. Z. & Palmer, A. G. III. $R_{1\rho}$ relaxation for two-site chemical exchange: general approximations and some exact solutions. *J. Magn. Reson.* **177**, 221–227 (2005).
46. Ferry, J. D., Grandine, L. D. & Fitzgerald, E. R. The relaxation distribution function of polyisobutylene in the transition from rubber-like to glass-like behavior. *J. Appl. Phys.* **24**, 911–916 (1953).
47. Denisov, V. P., Peters, J., Horlein, H. D. & Halle, B. Using buried water molecules to explore the energy landscape of proteins. *Nature Struct. Biol.* **3**, 505–509 (1996).
48. Olson, W. K. *et al.* A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* **313**, 229–237 (2001).
49. Abrescia, N. G., Gonzalez, C., Gouyette, C. & Subirana, J. A. X-ray and NMR studies of the DNA oligomer d(ATATAT): Hoogsteen base pairing in duplex DNA. *Biochemistry* **43**, 4092–4100 (2004).
50. Aishima, J. *et al.* A Hoogsteen base pair embedded in undistorted B-DNA. *Nucleic Acids Res.* **30**, 5244–5252 (2002).
51. MacKerell, A. D. Jr, Banavali, N. & Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **56**, 257–265 (2000).
52. Chocholoušová, J. & Feig, M. Implicit solvent simulations of DNA and DNA-protein complexes: agreement with explicit solvent vs experiment. *J. Phys. Chem. B* **110**, 17240–17251 (2006).
53. Feig, M. *et al.* Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures. *J. Comput. Chem.* **25**, 265–284 (2004).
54. Lee, M. S., Feig, M., Salsbury, F. R. Jr & Brooks, C. L. III. New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations. *J. Comput. Chem.* **24**, 1348–1356 (2003).
55. Nosé, S. A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys.* **81**, 511–519 (1984).
56. Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A* **31**, 1695–1697 (1985).
57. Frisch, M. J. *et al.* Gaussian 03, Revision C.02. (Gaussian, 2004).
58. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Submillimetre galaxies reside in dark matter haloes with masses greater than 3×10^{11} solar masses

Alexandre Amblard¹, Asantha Cooray^{1,2}, Paolo Serra¹, B. Altieri³, V. Arumugam⁴, H. Aussel⁵, A. Blain², J. Bock^{2,6}, A. Boselli⁷, V. Buat⁷, N. Castro-Rodríguez^{8,9}, A. Cava^{8,9}, P. Chanial¹⁰, E. Chapin¹¹, D. L. Clements¹⁰, A. Conley¹², L. Conversi³, C. D. Dowell^{2,6}, E. Dwek¹³, S. Eales¹⁴, D. Elbaz⁵, D. Farrah¹⁵, A. Franceschini¹⁶, W. Gear¹⁴, J. Glenn¹², M. Griffin¹⁴, M. Halpern¹¹, E. Hatziminaoglou¹⁷, E. Ibar¹⁸, K. Isaak¹⁴, R. J. Ivison^{18,4}, A. A. Khostovan¹, G. Lagache¹⁹, L. Levenson^{2,6}, N. Lu^{2,20}, S. Madden⁵, B. Maffei²¹, G. Mainetti¹⁶, L. Marchetti¹⁶, G. Marsden¹¹, K. Mitchell-Wynne¹, H. T. Nguyen^{6,2}, B. O'Halloran¹⁰, S. J. Oliver¹⁵, A. Omont²², M. J. Page²³, P. Panuzzo⁵, A. Papageorgiou¹⁴, C. P. Pearson^{23,24}, I. Pérez-Fournon^{8,9}, M. Pohlen¹⁴, N. Rangwala¹², I. G. Roseboom¹⁵, M. Rowan-Robinson¹⁰, M. Sánchez Portal³, B. Schulz^{2,20}, Douglas Scott¹¹, N. Seymour²³, D. L. Shupe^{2,20}, A. J. Smith¹⁵, J. A. Stevens²⁵, M. Symeonidis²³, M. Trichas¹⁰, K. Tugwell²⁶, M. Vaccari¹⁶, E. Valiante¹¹, I. Valtchanov³, J. D. Vieira², L. Vigroux²², L. Wang¹⁵, R. Ward¹⁵, G. Wright¹⁸, C. K. Xu^{2,20} & M. Zemcov^{2,6}

The extragalactic background light at far-infrared wavelengths^{1–3} comes from optically faint, dusty, star-forming galaxies in the Universe with star formation rates of a few hundred solar masses per year⁴. These faint, submillimetre galaxies are challenging to study individually because of the relatively poor spatial resolution of far-infrared telescopes^{5,6}. Instead, their average properties can be studied using statistics such as the angular power spectrum of the background intensity variations^{7–10}. A previous attempt¹¹ at measuring this power spectrum resulted in the suggestion that the clustering amplitude is below the level computed with a simple ansatz based on a halo model¹². Here we report excess clustering over the linear prediction at arcminute angular scales in the power spectrum of brightness fluctuations at 250, 350 and 500 μm . From this excess, we find that submillimetre galaxies are located in dark matter haloes with a minimum mass, M_{min} , such that $\log_{10}[M_{\text{min}}/M_{\odot}] = 11.5^{+0.7}_{-0.2}$ at 350 μm , where M_{\odot} is the solar mass. This minimum dark matter halo mass corresponds to the most efficient mass scale for star formation in the Universe¹³, and is lower than that predicted by semi-analytical models for galaxy formation¹⁴.

Despite recent successes in attributing most of the extragalactic background light at submillimetre wavelengths to known galaxy populations through stacking analyses^{15–17}, we have not individually detected the faint galaxies that are responsible for more than 85% of the total extragalactic intensity at these wavelengths¹⁸. The faint star-forming galaxies are expected to trace the large-scale structure of the Universe, especially in models where galaxy formation and evolution is closely connected to dark matter haloes. Although galaxies are not individually detected in low-resolution observations, their clustering is expected to leave a distinct signature in the total intensity variations at submillimetre wavelengths. The amplitude of the power spectrum of intensity variations as a function of the angular scale provides details on the redshift distribution and the dark-matter-halo mass scale of dusty, star-forming galaxies in the Universe⁷.

For this analysis, we use data from the Herschel Multi-tiered Extragalactic Survey¹⁸ (HerMES), recorded with the Spectral and Photometric Imaging Receiver (SPIRE¹⁹) on board the Herschel Space Observatory²⁰, during the Science Demonstration Phase of Herschel. The data come from a wide, 218'–by–218', area in the Lockman Hole overlapping with the Spitzer Wide-Area Infrared Extragalactic Survey (SWIRE), complemented by a narrow, but very deep (30 repeated scans), map of the Great Observatories Origins Deep Survey North (GOODS-N) field covering 30' by 30'. These fields have been very well studied at other wavelengths and they are known to have a low Galactic dust density, making it easier to distinguish the extragalactic component we wish to study. The observing time to complete each of the two fields was about 13.5 h, observing simultaneously at 250, 350 and 500 μm .

To limit the influence of a few bright galaxies on the measurement of the power spectrum, we remove galaxies brighter than 50 mJy in all three passbands by masking pixels in our maps with values larger than 50 mJy per beam, as well as the neighbouring pixels. We use the cross-power spectrum of two submaps as our estimate of the sky power spectrum to remove the contribution from the instrumental noise and alleviate potential systematic effects. We correct the raw cross-power spectra for the effects of the angular response function of the instrument and the transfer function of the map-making process. The angular response is established from a set of SPIRE observations targeting Neptune, which is a strong, point-like source for SPIRE, and involving a fine sampling of the beam with a total of 700 scans²¹. The effects of the filtering of the time-ordered data and of the map pixelization are captured with a large set of sky simulations. To estimate our uncertainties, we propagate the errors from the beam measurement, and the simulations provide us with the instrumental and sky variance. The quadratic sum of these errors constitutes our error estimate.

The measured angular power spectrum (Fig. 1a) contains contributions from spatial variations in the brightness of Galactic dust clouds (cirrus brightness) at large angular scales, the clustering of galaxies

¹Department of Physics & Astronomy, University of California, Irvine, California 92697, USA. ²California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA. ³Herschel Science Centre, European Space Astronomy Centre, Villanueva de la Cañada, 28691 Madrid, Spain. ⁴Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK. ⁵Laboratoire AIM-Paris-Saclay, CEA/DSM/Irfu-CNRS-Université Paris Diderot, CE-Saclay, pt courrier 131, F-91191 Gif-sur-Yvette, France. ⁶Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, California 91109, USA. ⁷Laboratoire d'Astrophysique de Marseille, OAMP, Université Aix-Marseille, CNRS, 38 rue Frédéric Joliot-Curie, 13388 Marseille cedex 13, France. ⁸Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain. ⁹Departamento de Astrofísica, Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain. ¹⁰Astrophysics Group, Imperial College London, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK. ¹¹Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia V6T 1Z1, Canada. ¹²Department of Astrophysical and Planetary Sciences, CASA 389-UCB, University of Colorado, Boulder, Colorado 80309, USA. ¹³Observational Cosmology Lab, Code 665, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA. ¹⁴Cardiff School of Physics and Astronomy, Cardiff University, Queens Buildings, The Parade, Cardiff CF24 3AA, UK. ¹⁵Astronomy Centre, Department of Physics & Astronomy, University of Sussex, Brighton BN1 9QH, UK. ¹⁶Dipartimento di Astronomia, Università di Padova, Vicolo Osservatorio, 3, 35122 Padova, Italy. ¹⁷ESO, Karl-Schwarzschild-Strasse 2, 85748 Garching bei München, Germany. ¹⁸UK Astronomy Technology Centre, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK. ¹⁹Institut d'Astrophysique Spatiale, Bâtiment 121, Université Paris-Sud 11 and CNRS (UMR 8617), 91405 Orsay, France. ²⁰Infrared Processing and Analysis Center, MS 100-22, California Institute of Technology, JPL, Pasadena, California 91125, USA. ²¹School of Physics and Astronomy, The University of Manchester, Alan Turing Building, Oxford Road, Manchester M13 9PL, UK. ²²Institut d'Astrophysique de Paris, UMR 7095, CNRS, UPMC Univ. Paris 06, 98bis Boulevard Arago, F-75014 Paris, France. ²³Space Science & Technology Department, Rutherford Appleton Laboratory, Chilton, Didcot OX11 0QX, UK. ²⁴Institute for Space Imaging Science, University of Lethbridge, Lethbridge, Alberta T1K 3M4, Canada. ²⁵Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK. ²⁶Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking RH5 6NT, UK.

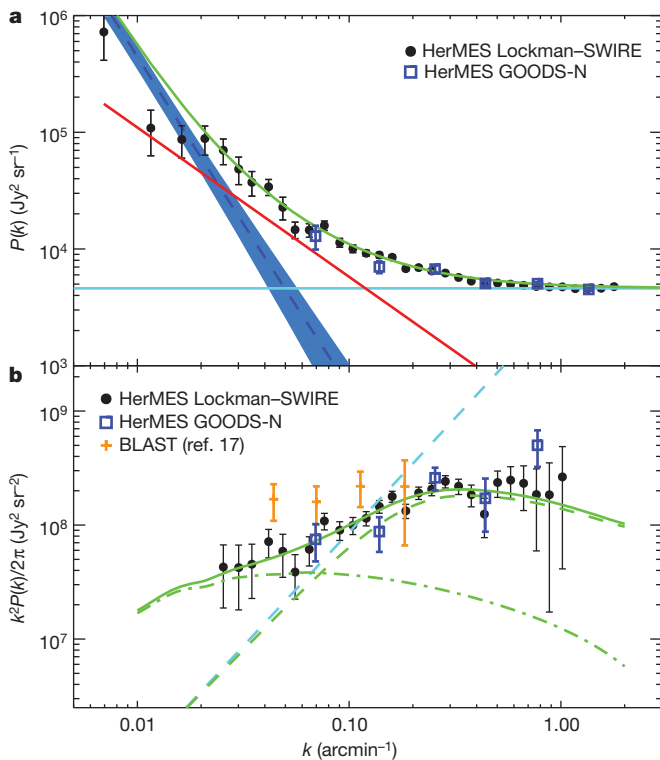


Figure 1 | The two-dimensional power spectrum of the Herschel map.

a, The total power spectrum, $P(k)$, at 350 μm as a function of the wavenumber, k , in inverse arcminutes. The error bars are uncertainties at the 68% confidence level. The shaded region is the cirrus signal in our Lockman-SWIRE field. To describe the total power spectrum, we take a power law with $P(k) = A(k/k_1)^n + P_{\text{SN}}$, where k_1 is fixed at 0.1 arcmin^{-1} and P_{SN} is the shot noise amplitude. At 350 μm , we find that $A = (5.79 \pm 0.26) \times 10^3 \text{ Jy}^2 \text{ sr}^{-1}$ and $n = -1.28 \pm 0.07$, shown by the red line. The light-blue line is the best-fit shot noise amplitude, which has a value of $4,600 \pm 70 \text{ Jy}^2 \text{ sr}^{-1}$, in agreement with the value of $4,500 \pm 220 \text{ Jy}^2 \text{ sr}^{-1}$ predicted by best-determined source counts²⁸. The shot noise errors include the 15% uncertainty in the SPIRE absolute flux calibration²¹. In green, we show the total power spectrum combined with the mean estimate of the cirrus signal. **b**, The angular clustering power spectrum at 350 μm as a function of wavenumber, with errors showing uncertainties at the 68% confidence level. The best-fit shot noise value (dashed blue line) has been removed from the data and its uncertainty added to the overall error in quadrature. The green lines show the best-fit halo model with a reduced χ^2 value of 1.02. The dot-dash line shows the two-halo term and the dashed line shows the one-halo term, which is responsible for the clustering at small angular scales. Data in the lowest-wavenumber bins contaminated by cirrus brightness have been omitted. For comparison, we also show a previous measurement of the power spectrum of brightness fluctuations at 350 μm , made with BLAST¹¹. The results related to 250 and 500 μm are summarized in Supplementary Information.

at intermediate angular scales and a white-noise component at small angular scales arising from the Poisson behaviour of the faint galaxies^{7,9}. The cirrus signal in our Lockman-SWIRE field is taken from existing measurements in the same field with the Infrared Astronomical Satellite, at 100 μm , and the Multiband Imaging Photometer for Spitzer²², and extends this spectrum from 100 μm to SPIRE wavelengths using the spectral dependence of a Galactic dust model²³. We remove this cirrus power spectrum from our measurements and account for the uncertainty in it by adding its error in quadrature to errors in our power spectrum points.

The Poisson behaviour of sources leads to an additional term in the angular power spectrum that is scale independent. The clustering component we measure is the excess of clustered background fluctuations above this shot noise level. As the confusion noise is at the level of 6 mJy at SPIRE wavelengths⁵, with fluctuations in the brightness of the background we are probing the clustering of faint galaxies with fluxes at the level of a few millijanskys at 350 μm . To extract astrophysical

information on faint galaxies from the clustering power spectrum, we make use of the halo model¹². This phenomenological approach connects the spatial distribution of galaxies in the Universe to that of dark matter. To model submillimetre galaxies in dark matter haloes, we use the following halo occupation distribution, which describes the number of galaxies as a function of the halo mass, M , when M is greater than a minimum mass scale, M_{min} : $N_{\text{gal}}(M) = 1 + (M/M_1)^\alpha$. Here M_1 is the mass scale at which more than one galaxy is present in a single dark matter halo, taken to be between 10 to 25 times M_{min} , and α is the power-law scaling of satellite galaxies with halo mass. The halo model involves two parts: the one-halo term, with clustering of galaxies within haloes, and the two-halo term, with clustering of galaxies between haloes. Whereas with the two-halo term alone, parameters related to the occupation number are degenerate with each other and the bias factor or the number density of galaxies, with clustering in the one-halo part of the power spectrum also included, the parameter degeneracies are broken and M_{min} can be determined more accurately¹².

At scales of a few arcminutes and above, we measure a clustering excess, arising from the one-halo term, over the two-halo term tracing the linear-density-field power spectrum scaled by galaxy bias (Fig. 1b). The one-halo term arises when more than one far-infrared galaxy occupies the same halo. A hint of the one-halo term was previously seen in the clustering of the bright (>30 -mJy) submillimetre galaxies²⁴, but evidence for such clustering was not found for bright galaxies in a different Herschel data set²⁵. To describe the power spectrum of the intensity fluctuations, we also need a prescription for the redshift evolution of the source intensity. Although models exist in the literature^{26,27}, our data are of sufficient quality that we can directly constrain the redshift evolution of the source intensity from our measurements: we constrain its value in four redshift bins in the range 0–4.0. The halo model parameters, the source intensity parameters and the shot noise amplitude are jointly estimated with Markov chain Monte Carlo fits to the power spectrum measurements. We impose the additional prior on our parameter estimates that the redshift-integrated source intensity from our model fits, including the fractional contribution from bright sources that we have masked, be within the 68% range of the known background light intensity in each of the three wavebands². We combine the estimates of the source intensity evolution in the three wavebands (250, 350 and 500 μm) and in the four redshift bins to measure the bolometric luminosity density between 8 and 1,100 μm as a function of the redshift (Fig. 2). We find that the far-infrared luminosity density continues to be significant out to a redshift of four and is at least a factor of ten larger than the luminosity density of individually detected submillimetre sources with flux densities above 30 mJy alone¹⁸.

Using the halo model fits, we estimate that the minimum dark matter mass scale for dusty, star-forming galaxies at the peak of the star formation history of the Universe is such that $\log_{10}[M_{\text{min}}/M_\odot] = 11.5^{+0.7}_{-0.2}$ at 350 μm , with a bias factor for the galaxies of $2.4^{+1.0}_{-0.2}$. For the minimum halo masses at 250 and 500 μm , $\log_{10}[M_{\text{min}}/M_\odot] = 11.1^{+1.0}_{-0.6}$ and $11.8^{+0.4}_{-0.3}$, respectively. The corresponding bias factors for the galaxies are $2.0^{+0.9}_{-0.1}$ and $2.8^{+0.4}_{-0.5}$. The differences in the minimum halo masses and the bias factors between the three wavelengths are probably due to a combination of effects including overall calibration uncertainties, the fact that at longer wavelengths we may be probing colder dust than at shorter wavelengths, and differences in the prior assumption on the total background intensity. In future, numerical models on the distribution of submillimetre galaxies will become useful in properly understanding some of these subtle differences. Averaged over the three wavelengths, the minimum halo mass for submillimetre galaxies is at the level of $3 \times 10^{11} M_\odot$, with an overall statistical uncertainty of roughly ± 0.4 in $\log_{10}[M_{\text{min}}/M_\odot]$.

On the basis of a variety of observed scaling relations such as that between stellar mass and circular velocity, the dark matter halo mass scale for efficient star formation has been indirectly inferred to be about $10^{11} M_\odot$ (ref. 13). As the submillimetre galaxies are the most active star-forming galaxies in the Universe, it is likely that the minimum halo

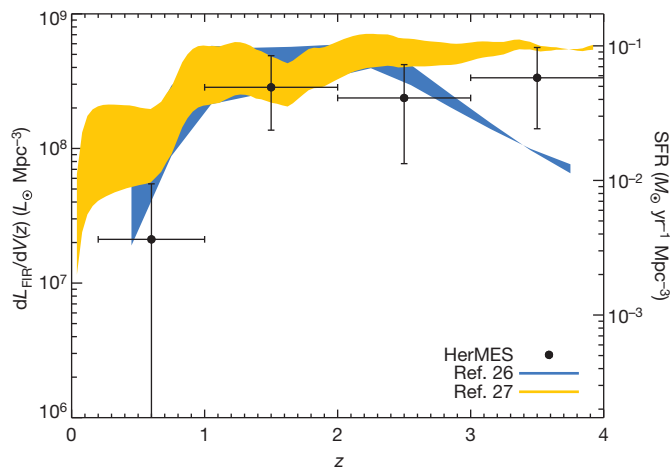


Figure 2 | Far-infrared bolometric luminosity density (8–1,110 μm) and star formation rate (SFR) as a function of redshift. The far-infrared bolometric luminosity density was estimated using the values of the source intensity in four redshift bins between redshifts $z = 0$ and 4 derived from model best fits to each of the power spectra from the three wavebands and with prior selection of models with $\alpha > 1$ for the power-law slope of the occupation number. We have assumed a modified black-body spectrum for the spectral energy distribution of the spatially unresolved sources, with an emissivity index of 1.5 and a dust temperature of 28 ± 8 K (ref. 29). We propagate the uncertainty in the dust temperature as an additional error when computing errors in the luminosity density in each of the redshift bins. The vertical bars show errors at the 68% confidence level propagated from errors in the source intensity evolution estimates and from the temperature prior, and the horizontal bars indicated the redshift range of each of the bins used for model fits. The yellow and blue shaded areas represent models^{26,27} with the same flux cut ($S > 50$ mJy) and the same temperature prior as our data. Our measurements are consistent with both these models. The redshift evolution of the far-infrared luminosity is a measure of the star formation history of the Universe³⁰, and by converting our estimate we find respective star formation rates of $< 0.01 M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}$, $(0.05 \pm 0.03) M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}$, $(0.04 \pm 0.03) M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}$ and $(0.06 \pm 0.04) M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}$ in the redshift bins with $z < 1$, $1 < z < 2$, $2 < z < 3$ and $3 < z < 4$. L_{\odot} , solar luminosity.

mass scale for such galaxies that we determine from brightness fluctuations corresponds to the preferred mass scale of active star formation in the Universe. In dark matter haloes below this mass, star formation is expected to be inefficient as a result of photoionization feedback¹³. The underlying astrophysics needed to explain the numerical value we find is still missing from galaxy formation theories: existing semi-analytical models predict a mass scale for faint, submillimetre galaxies that are roughly ten times larger¹⁴. We provide strong evidence for a minimum mass scale for active star-forming galaxies by studying the background intensity variations generated by those galaxies in our sky maps. Our direct estimate of the minimum dark matter halo mass provides a critical value needed to improve theoretical models of submillimetre galaxies and the overall picture of galaxy formation and evolution.

Received 11 August; accepted 20 December 2010.

Published online 16 February 2011.

1. Puget, J.-L. *et al.* Tentative detection of a cosmic far-infrared background with COBE. *Astron. Astrophys.* **308**, L5–L8 (1996).
2. Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L. & Shafer, R. A. The spectrum of the extragalactic far-infrared background from the COBE FIRAS observations. *Astrophys. J.* **508**, 123–128 (1998).
3. Dwek, E. *et al.* The COBE Diffuse Infrared Background Experiment Search for the cosmic infrared background. IV. Cosmological implications. *Astrophys. J.* **508**, 106–122 (1998).
4. Hughes, D. *et al.* High-redshift star formation in the Hubble Deep Field revealed by a submillimetre-wavelength survey. *Nature* **394**, 241–247 (1998).
5. Nguyen, H. T. *et al.* HerMES: the SPIRE confusion limit. *Astron. Astrophys.* **518**, L5 (2010).
6. Hauser, M. G. & Dwek, E. The cosmic infrared background: measurements and implications. *Annu. Rev. Astron. Astrophys.* **39**, 249–307 (2001).
7. Amblard, A. & Cooray, A. Anisotropy studies of the unresolved far-infrared background. *Astrophys. J.* **670**, 903–911 (2007).

8. Haiman, Z. & Knox, L. Correlations in the far-infrared background. *Astrophys. J.* **530**, 124–132 (2000).
9. Knox, L., Cooray, A., Eisenstein, D. & Haiman, Z. Probing early structure formation with far-infrared background correlations. *Astrophys. J.* **550**, 7–20 (2001).
10. Negrello, M. *et al.* Astrophysical and cosmological information from large-scale submillimetre surveys of extragalactic sources. *Mon. Not. R. Astron. Soc.* **377**, 1557–1568 (2007).
11. Viero, M. P. *et al.* BLAST: correlations in the cosmic far-infrared background at 250, 350, and 500 μm reveal clustering of star-forming galaxies. *Astrophys. J.* **707**, 1766–1778 (2009).
12. Cooray, A. & Sheth, R. Halo models of large scale structure. *Phys. Rep.* **372**, 1–129 (2002).
13. Bouché, N. *et al.* The impact of cold gas accretion above a mass floor on galaxy scaling relations. *Astrophys. J.* **718**, 1001–1018 (2010).
14. Gonzalez, J. E., Lacey, C. G., Baugh, C. M. & Frenk, C. S. The role of submillimetre galaxies in hierarchical galaxy formation. *Mon. Not. R. Astron. Soc.* (submitted); preprint at (<http://arxiv.org/abs/1006.0230>) (2010).
15. Devlin, J. M. *et al.* Over half of the far-infrared background light comes from galaxies at $z \geq 1.2$. *Nature* **458**, 737–739 (2009).
16. Dole, H. *et al.* The cosmic infrared background resolved by Spitzer. Contributions of mid-infrared galaxies to the far-infrared background. *Astron. Astrophys.* **451**, 417–429 (2006).
17. Marsden, G. *et al.* BLAST: resolving the cosmic submillimetre background. *Astrophys. J.* **707**, 1729–1739 (2009).
18. Oliver, S. *et al.* HerMES: SPIRE galaxy number counts at 250, 350 and 500 microns. *Astron. Astrophys.* **518**, L21 (2010).
19. Griffin, M. J. *et al.* The Herschel-SPIRE instrument and its in-flight performance. *Astron. Astrophys.* **518**, L3 (2010).
20. Pilbratt, G. *et al.* Herschel Space Observatory – an ESA facility for far-infrared and submillimetre astronomy. *Astron. Astrophys.* **518**, L1 (2010).
21. Swinyard, B. *et al.* In-flight calibration of the Herschel-SPIRE instrument. *Astron. Astrophys.* **518**, L4 (2010).
22. Lagache, G. *et al.* Correlated anisotropies in the cosmic far-infrared background detected by the multiband imaging photometer for Spitzer: constraint on the bias. *Astrophys. J.* **665**, L89–L92 (2007).
23. Schlegel, D. J., Finkbeiner, D. P. & Davis, M. Maps of dust infrared emission for use in estimation of reddening and cosmic microwave background radiation foregrounds. *Astrophys. J.* **500**, 525–534 (1998).
24. Cooray, A. HerMES: halo occupation number and bias properties of dusty galaxies from angular clustering measurements. *Astron. Astrophys.* **518**, L22 (2010).
25. Maddox, S. J. *et al.* Herschel ATLAS: the angular correlation function of submillimetre galaxies at high and low redshift. *Astron. Astrophys.* **518**, L11 (2010).
26. Lagache, G., Dole, H. & Puget, J.-L. Modelling the infrared galaxy evolution using a phenomenological approach. *Mon. Not. R. Astron. Soc.* **338**, 555–571 (2003).
27. Valiante, E. *et al.* A backward evolution model for infrared surveys: the role of AGN- and color- L_{TIR} distributions. *Astrophys. J.* **701**, 1814–1838 (2009).
28. Glenn, J. *et al.* HerMES: deep galaxy number counts from a $P(D)$ fluctuation analysis of SPIRE Science Demonstration Phase observations. *Mon. Not. R. Astron. Soc.* **409**, 109–121 (2010).
29. Amblard, A. *et al.* Herschel-ATLAS: dust temperature and redshift distribution of SPIRE and PACS detected sources using submillimetre colours. *Astron. Astrophys.* **518**, L9 (2010).
30. Kennicutt, R. C. Jr. Star formation in galaxies along the Hubble sequence. *Annu. Rev. Astron. Astrophys.* **36**, 189–232 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements SPIRE has been developed by a consortium of institutes led by Cardiff University (UK) and including the University of Lethbridge (Canada); NAOC (China); CEA and IAC (France); IFSI and the University of Padua (Italy); IAC (Spain); Stockholm Observatory (Sweden); Imperial College London, RAL, UCL-MSSL, UKATC and the University of Sussex (UK); and Caltech/JPL, IPAC and the University of Colorado (USA). This development has been supported by national funding agencies: CSA (Canada); NAOC (China); CEA, CNRS and CNRS (France); ASI (Italy); MCINN (Spain); SNSB (Sweden); STFC (UK); and NASA (USA). We thank M. Viero for comments. A.A., A. Cooray, P.S., A.A.K., K.M.-W. and other US co-authors are supported by NASA funds for US participants in Herschel through an award from JPL.

Author Contributions This paper represents the combined work of the HerMES collaboration, the SPIRE Instrument Team's Extragalactic Survey, and has been extensively internally reviewed. A. Cooray planned the study, supervised the research work of A.A. and P.S., and wrote the draft version of this paper. A.A. performed the power spectrum measurements and P.S. interpreted those measurements with the halo model. All other co-authors of this paper contributed extensively and equally by their varied contributions to the SPIRE instrument, Herschel mission, analysis of SPIRE and HerMES data, planning of HerMES observations and scientific support of HerMES, and by commenting on this manuscript as part of an internal review process.

Author Information The data presented in this paper are publicly available from the ESA/Herschel Science Archive (<http://herchel.esac.esa.int>) under the observational identifications 1342186108, 1342186109 and 1342185536. Derived products by the HerMES collaboration, such as source catalogues, will be released through HeDaM (<http://hedam.oamp.fr/HerMES>). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A. Cooray (acooray@ucl.ac.uk).

Geometric frustration in compositionally modulated ferroelectrics

Narayani Choudhury¹, Laura Walizer², Sergey Lisenkov³ & L. Bellaiche¹

Geometric frustration is a broad phenomenon that results from an intrinsic incompatibility between some fundamental interactions and the underlying lattice geometry^{1–7}. Geometric frustration gives rise to new fundamental phenomena and is known to yield intriguing effects such as the formation of exotic states like spin ice, spin liquids and spin glasses^{1–17}. It has also led to interesting findings of fractional charge quantization and magnetic monopoles^{5,6}. Mechanisms related to geometric frustration have been proposed to understand the origins of relaxor and multiferroic behaviour, colossal magneto-capacitive coupling, and unusual and novel mechanisms of high-transition-temperature superconductivity^{3–5,12,16}. Although geometric frustration has been particularly well studied in magnetic systems in the past 20 years or so, its manifestation in the important class formed by ferroelectric materials (which are compounds with electric rather than magnetic dipoles) is basically unknown. Here we show, using a technique based on first principles, that compositionally graded ferroelectrics possess the characteristic ‘fingerprints’ associated with geometric frustration. These systems have a highly degenerate energy surface and display critical phenomena. They further reveal exotic orderings with novel stripe phases involving complex spatial organization. These stripes display spiral states, topological defects and curvature. Compositionally graded ferroelectrics can thus be considered the ‘missing link’ that brings ferroelectrics into the broad category of materials able to exhibit geometric frustration. Our *ab initio* calculations allow deep microscopic insight into this novel geometrically frustrated system.

Geometrically frustrated systems such as spin ice and spin liquids reveal intriguing phenomena and are known to have a degenerate manifold of exotic ground states^{1–8}. Remarkable features of geometrically frustrated compounds include the formation of complex microstructures^{1–8}, the strong deviation of macroscopic properties from well-known critical behaviours^{9,10} and the presence of topological defects, spiral states and curvature^{11–17}. Here we report that compositionally modulated ferroelectrics exhibit all these phenomena.

We investigated compositionally graded (Ba,Sr)TiO₃ (BST) compounds, which are systems that are promising candidates for applications as storage capacitors and dielectrics for the next generation of dynamic random access memories^{18,19}. We chose the saw-like, compositional modulation that is depicted in Fig. 1a and that yields overall barium and strontium compositions of 50% each. The modulation is periodic along the *z* axis, which is taken to lie along the [001] pseudo-cubic direction (the *x* and *y* axes lie along the [100] and [010] directions, respectively). The mimicked modulated system is assumed to be epitaxially grown on a substrate and therefore to adopt the in-plane lattice constant of that substrate. The mis-fit strain arising from the difference between the in-plane lattice parameters of the compositionally graded material and the substrate is allowed to vary between –3% and +3%. (We note that epitaxial strain in oxide perovskite superlattices permits tuning of macroscopic properties and can also result in the formation of novel spatially ordered structures^{18–25}.) Monte Carlo simulations using effective Hamiltonians^{25,26} are performed to predict

properties of this strained modulated system for different periodic supercells (Methods Summary).

Figure 1b shows that the investigated material has a complex temperature/mis-fit strain phase diagram. At high temperature, the phase is the so-called paraelectric, *p*, state²² for all misfit strains, whereas a large tensile strain yields, at lower temperature, the so-called orthorhombic *aa* state²², which possesses an in-plane polarization lying along [110]. These two phases have also been found in BaTiO₃ thin films²³ and

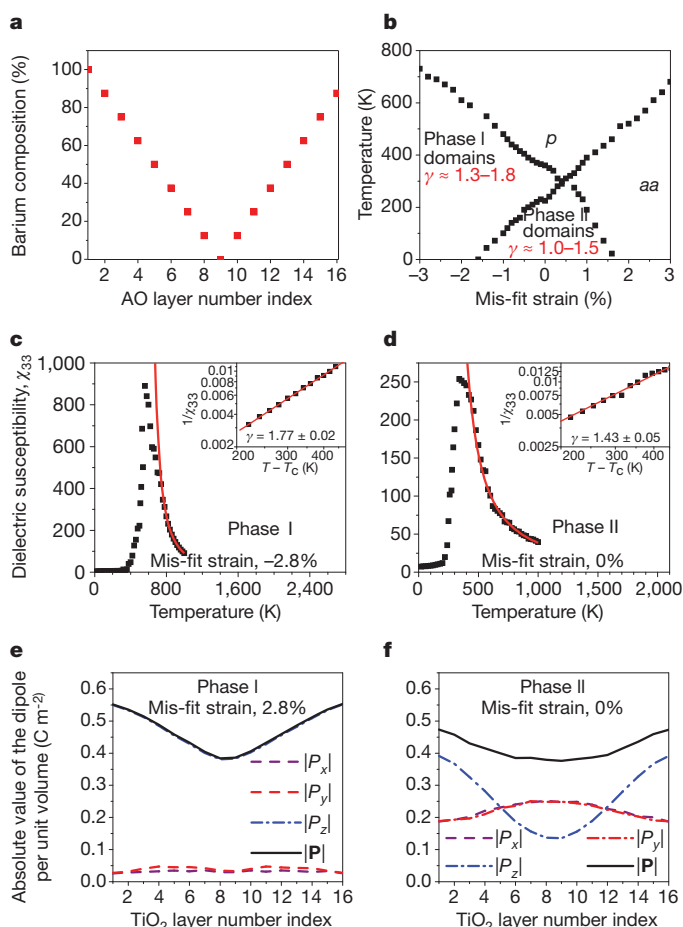


Figure 1 | Characteristics and material properties. **a**, Composition modulation along the [001] growth direction. **b**, Computed phase diagram. Typical error bars are 0.4% in mis-fit strain and 150 K in temperature. **c, d**, Temperature dependence of the static dielectric susceptibility and its log-log plot (inset) in phases I (**c**) and II (**d**). The subscripts ‘3’ refer to the [001] pseudo-cubic direction. The red solid lines represent fits of the dielectric susceptibility by the power law $1/\chi \propto (T - T_c)^\gamma$ for $T > T_c$. The possible ranges of γ for phases I and II are indicated in **b**. **e, f**, Layer-by-layer averages of the dipoles’ magnitude and Cartesian components in phases I (**e**) and II (**f**).

¹Department of Physics, University of Arkansas, Fayetteville, Arkansas 72701, USA. ²Engineer Research and Development Center, Vicksburg, Mississippi 39180, USA. ³Department of Physics, University of South Florida, Tampa, Florida 33620, USA.

BaTiO₃/SrTiO₃ superlattices²⁵. Similarly, the two other kinds of phase present in the phase diagram of Fig. 1b bear resemblance to some states discovered in some BaTiO₃/SrTiO₃ superlattices²⁵. These are ‘phase I’, which occurs for the largest compressive strains, below the Curie temperature, and which consists of separate ‘up’ and ‘down’ domains, for which the *z* components of the electric dipoles are respectively positive and negative; and ‘phase II’, which further has an homogeneous in-plane polarization along [110], in addition to forming out-of-plane domains, for intermediate strain and below a strain-dependent critical ferroelectric temperature, *T_c*. However, unlike in BaTiO₃/SrTiO₃ superlattices²⁵, the up and down domains in phases I and II do not ‘nicely’ organize into periodic nanostripes alternating along a specific, in-plane direction. The domain states forming phases I and II are not only exotic but are also found to be highly degenerate (see, for example, Fig. 2): various states with different complex organizations of up and down domains, but with similar energies, can be found for the same mis-fit strain (even for the largest studied supercell). Interestingly, anomalous microstructures and ground-state degeneracy are typical signatures of frustration^{1–7}.

The formation of the exotic structures of phases I and II is also associated with anomalous dielectric responses. More precisely, the dielectric susceptibility, χ , for temperatures, *T*, above the *p*-to-phase I and *p*-to-phase II transitions satisfies the power law $1/\chi \propto (T - T_c)^\gamma$ over a large range of temperatures, with the critical exponent, γ , significantly deviating from its Curie–Weiss value of one (Fig. 1c, d). In particular, γ can be as large as 1.8 in phase I, depending on the resulting microstructure. However, a similar analysis for the dielectric constant in the vicinity of the *p*-to-*aa* phase transition (that is, with no out-of-plane domains at low temperature and large tensile strain) yields $\gamma = 1$, in agreement with the Curie–Weiss behaviour. These dielectric anomalies in phases I and II are once again consistent with frustration, because geometric frustration has been documented to affect critical behaviours significantly^{9,10}. We further obtain a correlation between the exponent γ and the complexity of the microstructure: the larger γ is, the more the dipolar configuration differs from the ‘regular’, periodic nanostripe domains previously found in BaTiO₃ films²³ and BaTiO₃/SrTiO₃ superlattices²⁵. (We note that values of γ larger than one characterize a phase transition as being diffuse; a recent work linked diffuseness with dipolar inhomogeneities²⁰.) Such correlation is of great benefit in unravelling the microscopic initiation and development of frustration in the system we investigate. In fact, Fig. 2a (which corresponds to the less frustrated microstructure, as its associated γ parameter is the smallest among all

the displayed configurations) shows that frustration is initiated by the flipping of some up and down dipolar displacements near domain walls. The other panels of Fig. 2 demonstrate that frustration further percolates across to other sites, via flipping of dipoles, leading to very exotic shapes for the dipole pattern in the most frustrated systems—such as the ‘torus’ or ‘cross’ configurations in Fig. 2e, f. The torus configuration enables us to establish a parallel between our modulated system and geometrically frustrated spin liquids, which are known to have degeneracies on a torus^{1,8}. Again, such a parallel in behaviour strongly supports the idea that the ferroelectric material we study is also geometrically frustrated. This idea is further confirmed when realizing that difficulties in precisely predicting physical behaviour are a characteristic signature of geometric frustration¹, consistent with the fact that the numerous discovered possible ground states (see, for example, Fig. 2) span a wide range of critical exponents (see, for example, Fig. 1b).

As shown in Fig. 1e, f, a common feature emerges among the morphologically distinct states forming phases I and II: the average magnitude of the *z* component of the (titanium-centred) local electric dipoles varies smoothly as a function of the index of the (001) TiO₂ layer. This magnitude is maximal when the layer is close to a pure BaO plane and minimal when the TiO₂ layer is close to a pure SrO plane, consistent with the fact that decreasing the barium composition in BST solid solutions results in a smaller polarization at low temperature and to a smaller equilibrium lattice constant. We note that neither phase I nor phase II has an out-of-plane component of the overall polarization, as a result of the cancellation of the *z* component of the dipoles between the up and down domains.

It was also important to determine whether any non-trivial microscopic ordering between dipoles occurred in the frustrated states. Figure 3a shows a three-dimensional view of the dipolar arrangement of a ground state belonging to phase II, and Fig. 3b shows this dipole pattern in a given (010) plane, repeated twice along both the *x* and the *z* directions. We define a ‘single *z* chain’ as a chain oriented along the *z* direction and passing through all the (001) TiO₂ layers of the graded system. Figure 3c shows the average angle between dipoles belonging to the same *z* chain as a function of the distance between the centres of these dipoles. It can be seen from Fig. 3a, b that in most single *z* chains each dipole has either a positive or a negative *z* component (that is, the *z* components of the dipoles alter only in size from one (001) TiO₂ layer to another within a *z* chain). However, the in-plane components of the dipoles are found to significantly vary in direction when going

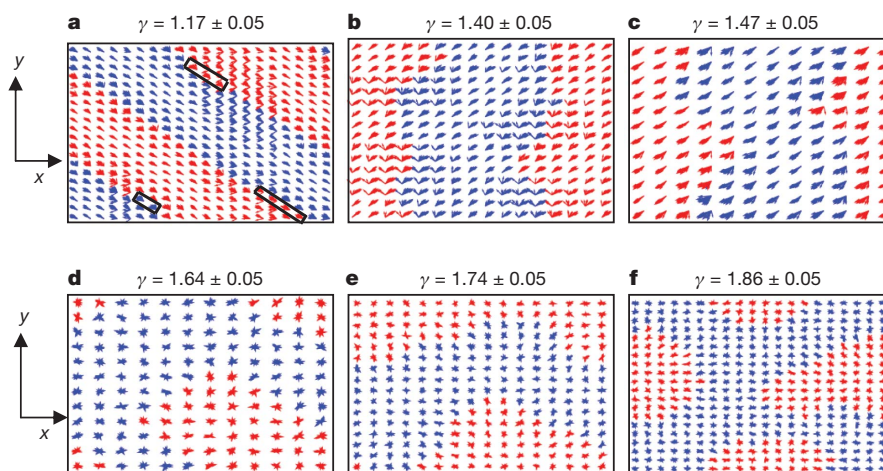


Figure 2 | Ground-state microstructures. Top views of various ground-state microstructures associated with phase II (a–c; mis-fit strain, 0%) and phase I (d–f; mis-fit strain, –2.8%) at *T* = 20 K. Blue, positive dipole *z* component; red, negative dipole *z* component. The exotic microstructures in a–c have similar internal energies, as do those in d–f. The critical exponent, γ , is indicated in each

panel. Selected flipped sites occurring near the domain walls are highlighted in a. Different supercell sizes are used for the same mis-fit strain (for instance, d, e and f correspond to $12 \times 12 \times 16$, $16 \times 16 \times 16$ and $20 \times 20 \times 16$ supercells, respectively).

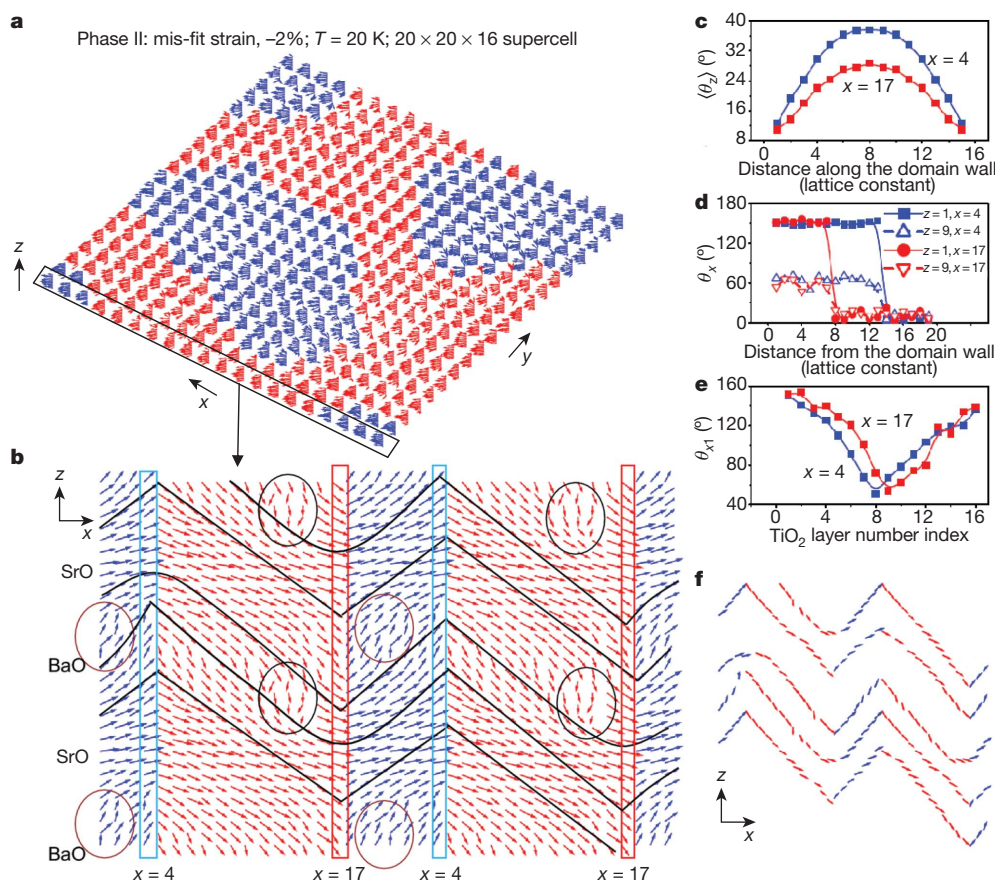


Figure 3 | Frustrated ground-state microstructure and dipolar orientational correlations in phase II. **a**, Ground-state dipolar configuration. Blue, positive dipole z component; red, negative dipole z component. **b**, Bending of lines, curvature and herringbone patterns in the microstructure are indicated by solid black lines. Selected z chains are highlighted and some

'topological defects' are circled. **c–e**, Orientational correlation functions (see text). For the blue curves the first dipole (used to compute these correlation functions) belongs to an up z chain, and for the red curves it belongs to a down z chain. **f**, Dipoles forming the herringbone pattern (black lines in **b**).

from one (001) TiO_2 layer to another along a given z chain. Such variation is reminiscent of the spiral or helical states that have been reported in other geometrically frustrated systems¹⁶. This leads to an average angle of $\sim 10^\circ$ between adjacent dipoles along any up (or down) z chain (Fig. 3c). Regarding the overall organization between z chains, Fig. 3a, b shows that within a given (001) TiO_2 layer, the dipoles can change the sign of their z components when scanning the different z chains along the x or y direction. Several down z chains (that is, in which the dipoles have negative z components) exist next to each other and then alternate with successive up z chains along the x or y direction. To appreciate such organization better, in Fig. 3d we display the correlated angle between a dipole belonging to a z chain and another dipole located along the same [100] pseudo-cubic direction as a function of the distance between the centres of these two dipoles, for two different (001) TiO_2 layers: one close to a pure BaO plane and one close to a pure SrO plane. Similarly, Fig. 3e shows the angle formed by two dipoles at the domain walls as a function of the index of the (001) TiO_2 layer. Figure 3b, d, e indicates that any (domain wall) connection between down and up z chains along the x or y axis involves a complex, cooperative dipolar arrangement. This is evidenced by the fact that the average angle between any two dipoles adjacent along the [100] pseudo-cubic direction and belonging respectively to up and down z chains strongly depends on the position of the (001) TiO_2 layer (Fig. 3e). For instance, this angle is around 150° for (001) TiO_2 layers close to pure BaO planes but is 60° for (001) TiO_2 layers close to pure SrO planes. Furthermore, away from the domain walls, neighbouring down (or up) z chains mostly possess dipoles that

are more-or-less constant in direction and magnitude within a given (001) TiO_2 layer (we note, however, that the 'topological defect' regions circled in Fig. 3b constitute an exception to this directional trend). Such features, in combination with the fact that the in-plane dipoles vary in direction when moving along z within a given z chain, result in complex curvatures in the domain patterns (see, for example, the curved lines in Fig. 3b, f; see Supplementary Fig. 1). It is important to note that curvature has been proposed as the driving mechanism of geometric frustration in cholesteric blue phases^{14,27}. Curvature leads to multiple frustrated orientational states. Cooperative phenomena and complex pattern formation are characteristic hallmarks of geometric frustration⁷. Cooperative dipolar arrangements also lead to herringbone-type domain patterns^{28–30} that are elongated along the x axis (Fig. 3b, f). The angle formed by the herringbone pattern's kinks depend on the position of the (001) TiO_2 layer (consistent with Fig. 3e). Each kink of the pattern lies on a domain wall. Herringbone patterns have also been reported in frustrated systems¹⁷.

Regarding the microscopic characteristics of phase I, Fig. 4 shows that the individual z chains can still have a variation in their in-plane dipole components along the growth direction (Fig. 4b), but to a much smaller extent than the z chains of phase II. This is a result of the fact that compressive strain acts to annihilate in-plane components of dipoles in favour of enhancing their z components²⁵. A striking feature of phase I is that the domain walls separating up domains from down domains can move between adjacent (010) planes (compare Fig. 4d with Fig. 4e, for example) and that the widths of the up and down domains can even change between neighbouring (010) planes (compare Fig. 4f with

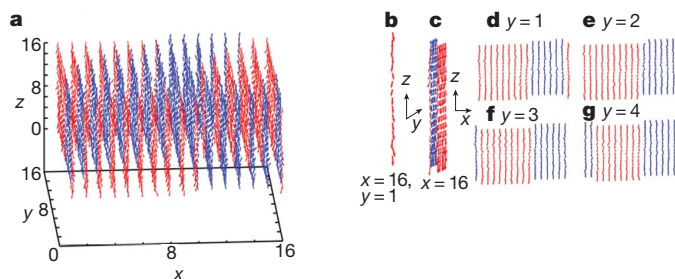
Phase I: mis-fit strain, -2.6% ; $T = 20$ K

Figure 4 | Frustrated ground-state microstructure and spiral states in phase I. **a**, Pattern of dipoles for a possible frustrated ground state. This three-dimensional view reveals spiral arrangements. **b**, Single z chain (see text). **c**, Assembly of z chains adjacent along the $[010]$ pseudo-cubic direction. **d–g**, Dipolar arrangements in different, successive (010) planes. Blue, positive dipole z component; red, negative dipole z component.

Fig. 4g, for example). The same features occur between adjacent (100) planes, hence leading to the propagation of two-dimensional waves (connecting the centres of gravity of neighbouring domain walls) inside the compositionally graded system. Such spatial modulation is reminiscent of that obtained in magnetism¹⁶, wherein frustration effects, when sufficiently strong, stabilize spiral states (see Supplementary Fig. 2). The dipolar arrangements reported here permit the domain walls to evolve through various representative domain configurations, all of which are equally probable, which explains the degeneracy of the exotic states associated with phase I. Moreover, the novel spatial arrangements of atoms depicted in Fig. 4 automatically lead to cancellation of both the out-of-plane and the in-plane components of the polarization in phase I; that is, the sum of all dipole vectors over the whole graded system vanishes.

A possible microscopic origin for geometric frustration in the compositionally modulated ferroelectric system studied here may reside in the fact that the z components of the dipoles in different (001) TiO_2 layers have different magnitudes, as a result of the change in barium composition along the $[001]$ pseudo-cubic direction (Fig. 1e, f). The difference in magnitude between adjacent (001) TiO_2 layers prevents the formation of states that are homogeneously polarized along $[001]$, because such formation would lead to energy-costly depolarizing fields (originating in the dipoles being inhomogeneous between different (001) TiO_2 layers). It also prevents the occurrence of the regular, periodic nanostripe domains found in BaTiO_3 films or long-period $\text{BaTiO}_3/\text{SrTiO}_3$ superlattices²⁵, because the difference in magnitude between the z components of the dipoles in any two adjacent (001) TiO_2 layers is costly in terms of short-range dipolar interactions (and thus adds to the cost of short-range energy associated with the formation of domain walls). The composition modulation along the growth direction therefore prevents the formation of regular polarized or domain states, because it dramatically affects long-range interactions that are related to depolarizing fields and short-range interactions. Such composition modulation thus naturally leads to frustrated systems, as long as the z components of the dipoles exist (consistent with the fact that we did not find any sign of frustration for large tensile strains in the corresponding *aa* ground state). Comparing between the different energetic terms of our effective-Hamiltonian approach for regular nanostripe domains and the reported microstructures also reveals that inhomogeneous strains have a key role in the formation of the microstructures and rich textures associated with phases I and II, in analogy with the finding of ref. 30 about the origins of some complex states. All the anomalous features described here (that are consistent with geometric frustration) were also numerically found in another BST modulated system. In that compound, the composition continuously changes by 25% along the growth direction (rather than by 12.5% as here; see Fig. 1a).

Our results thus strongly suggest that compositionally graded ferroelectric systems form a novel class of materials that can exhibit geometric frustration. They can therefore provide a new route to studying characteristics of geometric frustration, as compositionally graded ferroelectrics can be experimentally synthesized by, for example, molecular beam epitaxy and pulsed laser deposition²⁴. Future studies could also examine the influence of static and dynamic electric fields or defects on the behaviour of these modulated systems, which may be relevant for understanding the physics of relaxors and glasses and may have implications in diverse fields.

METHODS SUMMARY

The total internal energy, E_{tot} , of the studied BST graded system is provided by the effective-Hamiltonian approach of ref. 26, with parameters determined from first-principle calculations on relatively small supercells. This effective-Hamiltonian scheme has been shown to reproduce accurately the experimental composition-temperature phase diagram of disordered BST solid solutions²⁶, as well as to provide predictions that are in good agreement with *ab initio* calculations in $\text{BaTiO}_3/\text{SrTiO}_3$ superlattices²⁵. The degrees of freedom for E_{tot} are the homogeneous and inhomogeneous strains²⁶, and the local soft modes, \mathbf{u}_i , in all five-atom unit cells i . We note that \mathbf{u}_i is directly proportional to the electric dipole in cell i and is centred on the titanium sites, in contrast to the alloying that occurs on the A sites of the ABO_3 perovskite structure in BST. The barium and strontium atoms are randomly distributed in each (001) AO plane, with the overall barium and strontium compositions of each AO plane following the saw-like modulation depicted in Fig. 1a. The epitaxial growth of the investigated modulated BST material on a given substrate is mimicked by freezing some components of the homogeneous strain tensor^{22,25,26}, to impose the condition that each (001) layer has the same in-plane lattice constant as the chosen substrate²⁵. The zero in mis-fit strain corresponds to the predicted lattice constant of the cubic disordered $\text{Ba}_{0.5}\text{Sr}_{0.5}\text{TiO}_3$ alloy interpolated down to 0 K (ref. 26). The energy E_{tot} is used in Monte Carlo simulations (with up to 300,000 sweeps) on $12 \times 12 \times 16$, $16 \times 16 \times 16$ or $20 \times 20 \times 16$ supercells to obtain finite-temperature properties, such as the mis-fit strain phase diagram and static dielectric responses. The low-temperature microstructures are obtained by gradual cooling (in steps of 5 K) from 1,000 K, iteratively reading in configurations from simulated results for the previous temperature.

Received 29 June; accepted 9 December 2010.

Published online 9 February 2011.

- Moessner, R. & Ramirez, A. P. Geometric frustration. *Phys. Today* **59**, 24–29 (2006).
- Harris, M. J. Taking the frustration out of ice. *Nature* **399**, 311–312 (1999).
- Hemberger, J. *et al.* Relaxor ferroelectricity and colossal magnetocapacitive coupling in ferromagnetic CdCr_2S_4 . *Nature* **434**, 364–367 (2005).
- Anderson, P. W. The resonating valence bond state in La_2CuO_4 and superconductivity. *Science* **235**, 1196–1198 (1987).
- Laughlin, R. B. The relationship between high temperature superconductivity and the fractional quantum Hall effect. *Science* **242**, 525–533 (1988).
- Castelnovo, C., Moessner, R. & Sondhi, S. L. Magnetic monopoles in spin ice. *Nature* **451**, 42–45 (2008).
- Lee, S.-H. *et al.* Emergent excitations in a geometrically frustrated magnet. *Nature* **418**, 856–858 (2002).
- Wen, X. G. & Niu, Q. Ground state degeneracy of the fractional quantum Hall states in the presence of a random potential and on high genus Riemann surfaces. *Phys. Rev. B* **41**, 9377–9396 (1990).
- Hagemann, I. S., Khalifah, P. G., Ramirez, A. P. & Cava, R. J. Geometric magnetic frustration in olivines. *Phys. Rev. B* **62**, R771–R774 (2000).
- Lee, D. H., Joannopoulos, J. D., Negele, J. W. & Landau, D. P. Discrete-symmetry breaking and novel critical phenomena in an antiferromagnetic planar (XY) model in two dimensions. *Phys. Rev. Lett.* **52**, 433–436 (1984).
- Villain-Guillot, S., Dandolo, R., Saxena, A. & Bishop, A. R. Topological solitons and geometrical frustration. *Phys. Rev. B* **52**, 6712–6722 (1995).
- Mostovoy, M. Multiferroics go high- T_c . *Nature Mater.* **7**, 269–270 (2008).
- Libál, A., Reichhardt, C. J. O. & Reichhardt, C. Creating artificial ice states using vortices in nanostructured superconductors. *Phys. Rev. Lett.* **102**, 237004 (2009).
- Sethna, J. P. Frustration and curvature: glasses and the cholesteric blue phase. *Phys. Rev. Lett.* **51**, 2198–2201 (1983).
- Balents, L. Spin liquids in frustrated magnets. *Nature* **464**, 199–208 (2010).
- Cheong, S. W. & Mostovoy, M. Multiferroics: a magnetic twist for ferroelectricity. *Nature Mater.* **6**, 13–20 (2007).
- Papanikolaou, S., Raman, K. S. & Fradkin, E. Devil's staircases, quantum dimer models, and stripe formation in strong coupling models of quantum frustration. *Phys. Rev. B* **75**, 094406 (2007).
- Dawber, M., Rabe, K. M. & Scott, J. M. Physics of thin film ferroelectric oxides. *Rev. Mod. Phys.* **77**, 1083–1130 (2005).
- Okatan, M. B., Roytburd, A. L., Mantese, J. V. & Alpay, S. P. Domain engineering in compositionally graded ferroelectric films for enhanced dielectric response and tunability. *J. Appl. Phys.* **105**, 114106 (2009).

20. Bin-Omran, S., Kornev, I., Ponomareva, I. & Bellaiche, L. Diffuse phase transitions in ferroelectric thin films from first principles. *Phys. Rev. B* **81**, 094119 (2010).
21. Bousquet, E. *et al.* Improper ferroelectricity in perovskite oxide artificial superlattices. *Nature* **452**, 732–737 (2008).
22. Pertsev, N. A., Zembilgotov, A. G. & Tagantsev, A. K. Effect of mechanical boundary conditions on phase diagrams of epitaxial ferroelectric thin films. *Phys. Rev. Lett.* **80**, 1988–1991 (1998).
23. Lai, B.-K., Kornev, I., Bellaiche, L. & Salamo, G. Phase diagrams of epitaxial BaTiO₃ ultrathin films from first principles. *Appl. Phys. Lett.* **86**, 132904 (2005).
24. Lee, H. N., Christen, H. M., Chisholm, M. F., Rouleau, C. M. & Lowndes, C. H. Strong polarization enhancement in asymmetric three-component ferroelectric superlattices. *Nature* **433**, 395–399 (2004).
25. Lisenkov, S. & Bellaiche, L. Phase diagrams of BaTiO₃/SrTiO₃ superlattices from first principles. *Phys. Rev. B* **76**, 020102(R) (2007).
26. Walizer, L., Lisenkov, S. & Bellaiche, L. Finite-temperature properties of (Ba,Sr)TiO₃ systems from atomistic simulations. *Phys. Rev. B* **73**, 144105 (2006).
27. Sethna, J. P. Frustration, curvature and defect lines in metallic glasses and the cholesteric blue phase. *Phys. Rev. B* **31**, 6278–6297 (1985).
28. Vedmedenko, E. Y. *Competing Interactions in the Nanoworld* 3–16 (Wiley-VCH, 2007).
29. Narasimhan, S. & Vanderbilt, D. Elastic stress domains and the herringbone reconstruction on Au(111). *Phys. Rev. Lett.* **69**, 2455–2458 (1992).
30. Vanderbilt, D. in *Computations for the Nano-Scale* (eds Blöchl, P. E., Fisher, A. J. & Joachim, C.) 1–11 (Kluwer, 1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the US National Science Foundation, Office of Naval Research and Department of Energy. We gratefully acknowledge extensive use of the supercomputing resources of the University of Arkansas High Performance Computing Center as well as the Center for Piezoelectrics by Design, College of William and Mary, Virginia. We thank A. Apon, D. Chaffin, J. Pummill and E. J. Walter for computational support.

Author Contributions This work is an outgrowth of an ongoing project on compositionally modulated ferroelectrics at the University of Arkansas. L.W., S.L. and L.B. developed an effective-Hamiltonian implementation for BST systems. N.C. carried out the present Monte Carlo simulations using these effective-Hamiltonian and code implementations. N.C. found exotic degenerate ground states and spiral domains and suggested that these complex results can be explained in terms of geometric frustration. L.B. proposed further studies of critical behaviours and size dependency, and these additional simulations were carried out by N.C. Various complex details were jointly analysed by N.C. and L.B. and together they wrote the paper, with feedback from L.W. and S.L.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to N.C. (narayani@uark.edu).

Extended megadroughts in the southwestern United States during Pleistocene interglacials

Peter J. Fawcett¹, Josef P. Werne^{2,4,5}, R. Scott Anderson^{6,7}, Jeffrey M. Heikoop⁸, Erik T. Brown³, Melissa A. Berke³, Susan J. Smith⁷, Fraser Goff¹, Linda Donohoo-Hurley¹, Luz M. Cisneros-Dozal⁸, Stefan Schouten⁹, Jaap S. Sinninghe Damsté⁹, Yongsong Huang¹⁰, Jaime Toney⁸, Julianna Fessenden⁶, Giday WoldeGabriel⁶, Viorel Atudorei¹, John W. Geissman¹ & Craig D. Allen¹¹

The potential for increased drought frequency and severity linked to anthropogenic climate change in the semi-arid regions of the southwestern United States (US) is a serious concern¹. Multi-year droughts during the instrumental period² and decadal-length droughts of the past two millennia^{1,3} were shorter and climatically different from the future permanent, 'dust-bowl-like' megadrought conditions, lasting decades to a century, that are predicted as a consequence of warming⁴. So far, it has been unclear whether or not such megadroughts occurred in the southwestern US, and, if so, with what regularity and intensity. Here we show that periods of aridity lasting centuries to millennia occurred in the southwestern US during mid-Pleistocene interglacials. Using molecular palaeotemperature proxies⁵ to reconstruct the mean annual temperature (MAT) in mid-Pleistocene lacustrine sediment from the Valles Caldera, New Mexico, we found that the driest conditions occurred during the warmest phases of interglacials, when the MAT was comparable to or higher than the modern MAT. A collapse of drought-tolerant C₄ plant communities during these warm, dry intervals indicates a significant reduction in summer precipitation, possibly in response to a poleward migration of the subtropical dry zone. Three MAT cycles ~2 °C in amplitude occurred within Marine Isotope Stage (MIS) 11 and seem to correspond to the

mutated precessional cycles within this interglacial. In comparison with MIS 11, MIS 13 experienced higher precessional-cycle amplitudes, larger variations in MAT (4–6 °C) and a longer period of extended warmth, suggesting that local insolation variations were important to interglacial climatic variability in the southwestern US. Comparison of the early MIS 11 climate record with the Holocene record shows many similarities and implies that, in the absence of anthropogenic forcing, the region should be entering a cooler and wetter phase.

The hydroclimatology of the southwestern US shows significant natural variability including major historical droughts¹. Models of climate response to anthropogenic warming predict future dust-bowl-like conditions that will last much longer than historical droughts and have a different underlying cause, a poleward expansion of the subtropical dry zones⁴. At present, no palaeoclimatic analogues are available to assess the potential duration of aridity under a warmer climate or to evaluate its effect on the seasonality of precipitation.

Here we present a high-resolution climate record from an 82-m lacustrine sediment core (VC-3) from the Valles Caldera (Fig. 1) that spans two mid-Pleistocene glacial cycles from MIS 14 to MIS 10 (552 kyr ago to ~368 kyr ago; see Supplementary Information). MISs 11 and 13 are long interglacials that may have been as warm as

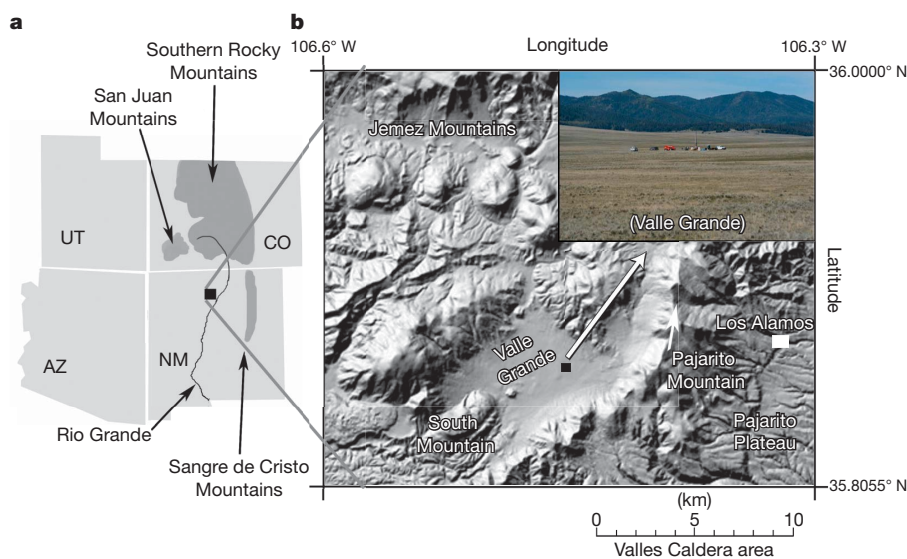


Figure 1 | Location map of the Valles Caldera. **a**, Location in northern New Mexico. **b**, Digital elevation model of the Valles Caldera showing the location of South Mountain rhyolite, Valle Grande, the drilling location of core VC-3 (black square) and a photograph of the drilling site.

¹Department of Earth & Planetary Sciences, University of New Mexico, Albuquerque, New Mexico 87131, USA. ²Large Lakes Observatory and Department of Chemistry and Biochemistry, University of Minnesota Duluth, Duluth, Minnesota 55812, USA. ³Large Lakes Observatory and Department of Geological Sciences, University of Minnesota Duluth, Duluth, Minnesota 55812, USA. ⁴Centre for Water Research, University of Western Australia, Crawley, Western Australia 6009, Australia. ⁵WA-Organic and Isotope Geochemistry Centre, Curtin University of Technology, Bentley, Western Australia 6845, Australia. ⁶School of Earth Sciences and Environmental Sustainability, Northern Arizona University, Flagstaff, Arizona 86011, USA. ⁷Laboratory of Paleocology, Bilby Research Center, Northern Arizona University, Flagstaff, Arizona 86011, USA. ⁸Earth and Environmental Sciences Division, EES-14, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA. ⁹NIOZ Royal Netherlands Institute for Sea Research, Department of Marine Organic Biogeochemistry, PO Box 59, 1790 AB Den Burg, Netherlands. ¹⁰Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA. ¹¹USGS Fort Collins Science Center, Jemez Mountains Field Station, Los Alamos, New Mexico 87544, USA.

the Holocene epoch, and MIS 11 is a good analogue for future natural climate variability with similar, low-amplitude precessional cycles^{6,7}. We used novel organic geochemical proxies (the cyclization ratio of branched tetraethers (CBT, related to pH) and the methylation index of branched tetraethers (MBT, related to temperature and pH^{5,8})) to reconstruct the annual MAT of the Valles Caldera watershed, and compared these with proxies of hydrologic balance to evaluate the relationship between warmth and aridity.

Interglacial MATs in the VC-3 record range from ~0 to 7 °C, with the highest temperatures occurring in MIS 13 and early in MIS 11 (Fig. 2). The highest temperatures (5–7 °C) are similar to modern MATs, of ~5 °C. The glacial stages have multiple millennial-scale temperature oscillations with amplitudes as large as 7 °C; approximately seven oscillations are preserved in MIS 12 (B1–B7), three in late MIS 14 (C1–C3) and one in early MIS 10 (A1). The frequency of these oscillations (2–10 kyr) is similar to those recorded in contemporaneous Atlantic Ocean sediment records⁹. All VC-3 stadials correlate with high percentages of *Picea* + *Abies* pollen, whereas interstadials have lower *Picea* + *Abies* pollen percentages and many correlate with local maxima in *Juniperus* and *Quercus* (Fig. 2). Increased percentages of Cyperaceae (sedge) pollen during several interstadials suggest a shallower lake rimmed by a broad marshy zone, which would have been minimized during stadials, when the lake was deeper. Interstadial shallowing probably resulted from increased evaporation and/or a reduction in the winter precipitation that dominates regional glacial-stage precipitation¹⁰.

Glacial terminations VI and V in the VC-3 record show temperature increases of ~7 and ~8 °C, respectively. The $\delta^{13}\text{C}$ record of TOC (Fig. 2) shows negative isotopic shifts of 2.5–3.5‰ at the terminations that we interpret as biotic responses to global increases in atmospheric CO_2 , similar to the Termination I $\delta^{13}\text{C}$ response in Lake Baikal¹¹.

We subdivide MIS 11 into five distinct substages, three warm and two cool, on the basis of MAT estimates, warm (lower-elevation) versus boreal (higher-elevation) pollen taxa, and variation in aquatic productivity proxies (Fig. 2). The warm substages (MISs 11a, c and e) are separated by intervals in which the temperature is ~2 °C lower (MISs 11b and d). Although these small temperature variations are within the error limits of the MBT/CBT calibration, their timing is supported by decreases in warm pollen taxa and increases in boreal pollen taxa (with the exception of MIS 11a). The warmest substage, MIS 11e, occurs early in the interglacial, and has peak MATs of 6–7 °C and the highest percentages of *Juniperus* pollen. After MIS 11e, the warm substages become progressively cooler.

The preservation of five MIS 11 substages in VC-3 is unusual. Most published records recognize only three substages, although a weak MIS 11e was noted in the Lake Baikal biogenic silica record¹² and there are three distinct (warm) peaks in MIS 11 pollen influx from Greenland preserved in ODP Site 646 sediments¹³. The VC-3 MIS 11a substage is cooler than the extended warm phases of MIS 11, similar to other mid-Pleistocene climate records¹², and is defined mainly by elevated lacustrine productivity (Si/Ti and TOC), more-positive $\delta^{13}\text{C}$ values, slightly higher temperature estimates and a combination of *Quercus*, *Picea* and *Abies* pollen that may not have a good modern climatic analogue. Within the limits of the VC-3 age model and the calibration uncertainty in the MBT/CBT proxies, the warm substages seem to correspond to the three precessional peaks of MIS 11, suggesting that the temperature response of this region to low-amplitude precessional cycles is ~2 °C. On the basis of the MIS 11 orbital forcing similarity with the Holocene, we suggest that in the absence of anthropogenic forcing future southwestern US climate should see a cooling of ~2 °C relative to the early Holocene.

Large parts of MIS 13 seem to have been warmer than most of MIS 11, as shown by MATs of up to 7 °C, higher *Juniperus* pollen percentages and the absence of *Picea* + *Abies* pollen (Fig. 2). Only MIS 11e had temperatures approaching the peak warmth of MIS 13. Other Northern Hemisphere records suggest that MIS 13 was warmer than

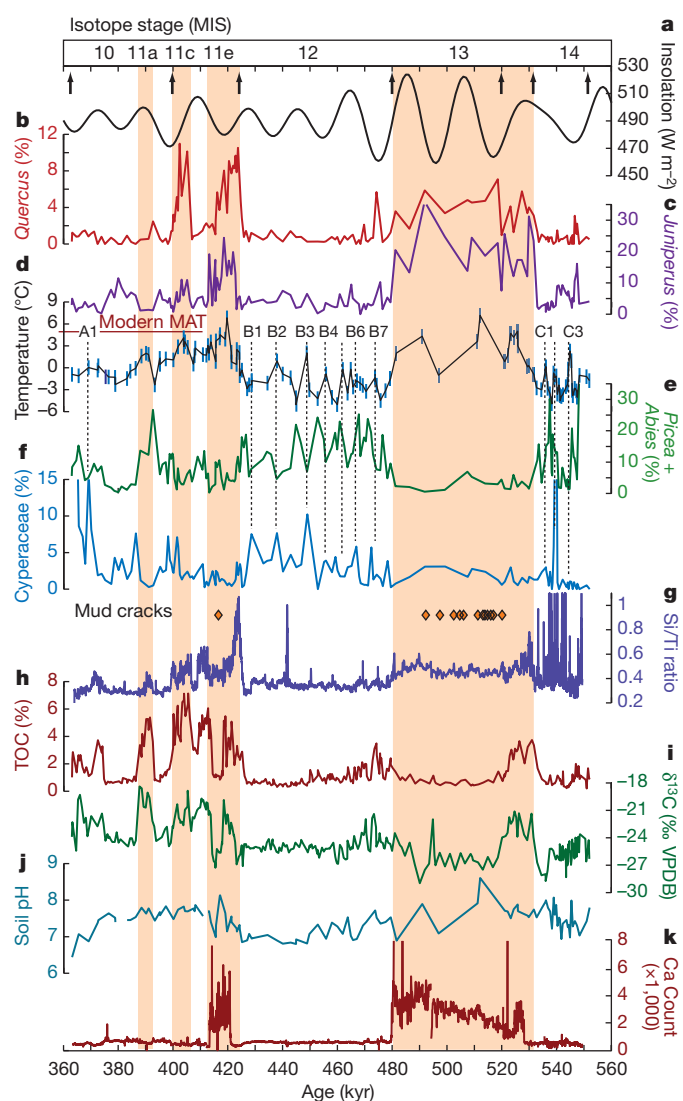


Figure 2 | Multi-proxy profiles of VC-3 plotted versus calendar age. Age model age–depth tie points are shown as arrows at the top and possible sedimentary hiatuses are indicated by positions of mud cracks (below). Shading indicates interglacial periods including odd-numbered MISs 11 and 13 and substages within MIS 11 (a, c and e). **a**, June insolation at latitude 30° N (ref. 6). **b**, *Quercus* (warm) pollen percentages. **c**, *Juniperus* (warm) pollen percentages. **d**, MAT estimates from MBT/CBT, with data size marker equivalent to 2 °C (blue). Red line shows the modern MAT, of 4.8 °C, in the Valle Grande. MBT/CBT temperature estimates have an absolute uncertainty of 5 °C based on uncertainties in the global calibration (see further discussion in Supplementary Information). Millennial-scale events within the three glacial periods, defined by local maxima in MAT and Cyperaceae and local minima in *Picea* and *Abies*, are indicated (A for MIS 10, B for MIS 12 and C for MIS 14). **e**, *Picea* + *Abies* (boreal) pollen percentages. **f**, Cyperaceae pollen percentages; mud cracks indicated with orange diamonds. **g**, Si/Ti ratios from core scanning X-ray fluorescence (XRF). Large peaks in MIS 14 correspond to pumiceous gravels. **h**, Total organic carbon (TOC). **i**, $\delta^{13}\text{C}_{\text{TOC}} = ((^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{standard}} - 1) \times 1,000\text{‰}$, relative to the Vienna Pee Dee Belemnite (VPDB) standard. **j**, Watershed soil pH estimate from CBT. **k**, Calcium concentration in sediments from core scanning XRF.

MIS 11^{14,15}, and a smaller Greenland ice sheet¹³ and a lack of ice rafting in the North Atlantic⁹ also indicate Northern Hemisphere warmth during MIS 13, although not necessarily more than during MIS 11. In contrast, Southern Hemisphere records uniformly show a cooler MIS 13^{14,16}.

The higher MIS 13 temperatures in the southwestern US occur despite lower interglacial values of atmospheric CO_2 and CH_4 (ref. 17). However, the amplitude of precessional cycles and, hence, extremes in

Northern Hemisphere insolation were larger in MIS 13 than MIS 11⁶ and may have led to higher continental temperatures during parts of MIS 13. Temperature variability during MIS 13 was as much as 6 °C, which is significantly larger than that during MIS 11 (Fig. 2). In combination with the apparent precessional timing of MIS 11 warm sub-stages, this suggests that the southwestern US responded more strongly to insolation variations than to interglacial trends in greenhouse gases or global ice volume.

Mud cracks present in the warmest phases, MIS 11e and MIS 13, are unambiguous indicators of drier conditions. One 70-cm mud crack occurs within MIS 11e, and ~3 m of section within the upper portion of MIS 13 sediments contains multiple, centimetre-scale mud cracks, making this portion of the VC-3 age model less certain (Fig. 2). Also, the presence or absence of calcite in VC-3 sediments provides a continuous indicator of closed-basin or, respectively, open-basin conditions in the lake. No calcite precipitated during freshwater open-basin conditions, whereas during drier (closed-basin) conditions, evaporative concentration led to calcite precipitation and preservation. XRF core scanning results show two intervals with high calcium concentrations during MISs 11e and MIS 13 (Fig. 2) that correlate with elevated (1–2%) total inorganic carbon (not shown), whereas core sections with low calcium concentrations have essentially no total inorganic carbon. Significant increases in calcium mark the onsets of closed-basin conditions coincident with rapid temperature increases a few thousand years after Terminations V and VI (Fig. 2). Mud cracks develop later within these closed-basin periods.

Long-term changes in watershed hydrology are also reflected in the CBT-derived soil pH record. Changes in soil pH are assumed to reflect changes in total precipitation¹⁸; greater soil leaching and acidification occurs with more precipitation, whereas drier conditions result in weaker soil acidification. The most alkaline soils occur within the interglacial mud-cracked facies (Fig. 2) and are more basic for longer periods in MIS 13. In contrast, soil pH shows a progressive acidification through MIS 12, consistent with progressively wetter conditions through that glacial stage, and possibly also caused by increases in boreal tree vegetation.

During MIS 11e, Si/Ti (a proxy for diatom productivity) is initially very high and then declines to average interglacial values, whereas TOC increases to ~5% in MIS 11e and is also high during early MIS 11d and MISs 11c and 11a. After the glacial termination, $\delta^{13}\text{C}_{\text{TOC}}$ rapidly increases to ~–20‰, indicating an expansion of C_4 plants in the watershed and higher lacustrine productivity levels¹⁹. Increases $\delta^{13}\text{C}_{\text{TOC}}$ also occurs in MISs 11c and 11a, and early in MIS 11d. Continued high Si/Ti, TOC and more-positive $\delta^{13}\text{C}_{\text{TOC}}$ values during early stages of closed-basin conditions in MIS 11e and MIS 13 indicate periods of robust summer precipitation and productivity related to insolation forcing of monsoon strength, even as reduced winter precipitation led to less precipitation overall and closed-basin conditions.

Mud-cracked facies in MIS 11e and MIS 13 are characterized by negative shifts in $\delta^{13}\text{C}_{\text{TOC}}$ of 6–7‰ and dramatic decreases in percentage TOC (Fig. 2). Si/Ti ratios, however, remain elevated relative to glacial values, suggesting that low percentage TOC values are due to organic degradation in shallow, oxidized sediments rather than lower aquatic productivity. The large negative shifts in $\delta^{13}\text{C}_{\text{TOC}}$ are best explained by a collapse of the interglacial C_4 plant community. Variations in C_3 and C_4 plant communities are a complex function of temperature, atmospheric CO_2 , and growing-season precipitation^{20,21}. These dry intervals include some of the highest MATs in the VC-3 record that should favour C_4 plants, and the relatively high interglacial levels of atmospheric CO_2 during MISs 11 and 13 vary by less than 20–30 p.p.m.v. Thus, the best explanation for the decline of C_4 plants in the watershed is a significant decrease in summer precipitation. In contrast to the early interglacial closed-basin phases where significant C_4 plant growth provided evidence for robust summer precipitation, we interpret the extended arid periods later in MIS 11e and MIS 13 to be the result of greatly reduced summer precipitation.

Following the aridity of MIS 11e, the lake expanded during MIS 11d as shown by well-laminated sediments and open-basin conditions (low calcium values). Despite this interval being ~2 °C cooler, sufficient summer rainfall early in MIS 11d allowed renewed C_4 plant growth.

Northern New Mexico at present receives ~40–50% of its annual precipitation total during the summer monsoon²². During the warmest phases of the interglacials, we would expect greater summer precipitation, as the monsoon is primarily driven by land surface heating²². Indeed, linkages among interglacial warmth, robust summer precipitation and precessional variations are indicated by the presence of C_4 plants in early MIS 13, early MIS 11e and MISs 11c and 11a, when MATs were similar to or slightly less than modern values, but the warmest intervals did not have robust summer precipitation. As possible analogues for interglacial aridity, both historical droughts and pre-historical megadroughts were characterized by reductions in winter precipitation as a consequence of more-frequent La Niña events^{2,3,23}, with summer precipitation reduced also.

In contrast, the extended arid episodes (centuries to millennia) of MIS 11e and MIS 13 lasted much longer than pre-historical megadroughts. An analogous relationship between peak interglacial warmth and extended aridity was also noted in a mid-Holocene bog record from the margin of the Valles Caldera²⁴. Here, ~2 kyr of desiccation occurred contemporaneously with the highest temperatures of the Holocene in the southwestern US²⁵ and with the northernmost extent of the inter-tropical convergence zone in the Gulf of Mexico²⁶. The timing of this dry episode in the Holocene interglacial following the deglaciation is very similar to that of the arid episode in MIS 11e; subsequent late-Holocene conditions became wetter in the southwestern US, with increased winter precipitation²⁷ similar in timing to wetter conditions during MIS 11d in the VC-3 record.

The strong correspondence between the warmest temperatures and extended aridity during at least three interglacials (MIS 13, MIS 11e and the early Holocene) in the southwestern US suggests a stable climate state fundamentally different from conventional drought conditions. These periods of aridity are related to lower winter precipitation (as mid-latitude westerlies shifted polewards during warmer periods), but reductions in summer precipitation seem to be critical to their development. Unlike the temporary summer blocking high over the southwestern US thought to partly explain the 1950s drought²⁸, these longer periods of aridity indicate a more permanent change in atmospheric circulation. Climate model analysis shows that the dust-bowl-like conditions predicted for the southwestern US over the next century in response to anthropogenic warming arise from a poleward shift of the mid-latitude westerlies and the poleward branch of the Hadley cell⁴. This response to warming is not transient and would result in a more arid southwestern US as long as the underlying conditions (warming) remained in place. Our palaeoclimate record shows that extended interglacial aridity is strongly linked to higher-than-modern temperatures and reduced summer rainfall, and we suggest that a similar expansion of the subtropical dry zone has occurred several times in the past in response to natural warming, even though MIS 11 and MIS 13 had different orbital and atmospheric CO_2 forcings. Our results strongly indicate that interglacial climates in the southwestern US can experience prolonged periods of aridity, lasting centuries to millennia, with profound effects on water availability and ecosystem composition. The risk of prolonged aridity is likely to be heightened by anthropogenic forcing^{1,4}.

METHODS SUMMARY

Measurement of fossil branched glycerol dialkyl glycerol tetraether (GDGT) membrane lipids from soil bacteria were conducted at the Royal Netherlands Institute for Sea Research (NIOZ) and Brown University following procedures outlined in Supplementary Information. At NIOZ we analysed GDGTs on an Agilent 1100 series LC-MSD SL, and at Brown University we analysed GDGTs on an HP 1200 series LC-MS. Both labs used an Alltech Prevail Cyano column (2.1 × 150 mm, 3 µm) with the same solvent elution scheme and instrument operating conditions. GDGTs were detected using atmospheric-pressure chemical ionization mass spectrometry. All

liquid chromatography/mass spectrometry runs were integrated at NIOZ by the same technician to ensure consistency. To evaluate the compatibility between the Brown and NIOZ measurements, representative samples were analysed on both machines and the resulting MBT/CBT indices were found to be identical within analytical uncertainty.

Processing for pollen included suspension in KOH, dilute HCL, hydrofluoric acid and acetolysis solution. The pollen sum included all terrestrial pollen types; Cyperaceae percentages were calculated outside the sum. We identified pollen grains to the lowest taxonomic level using the modern pollen reference collection at Northern Arizona University. Analysis for organic carbon elemental concentrations and $\delta^{13}\text{C}_{\text{TOC}}$ included samples being dried, ground and pretreated twice with 6 N HCL at 60 °C to remove the carbonate fraction. TOC and $\delta^{13}\text{C}_{\text{TOC}}$ were analysed using a Costech Elemental Analyser coupled to a Thermo-Finnigan Delta Plus isotope ratio mass spectrometer. The bulk elemental composition of core VC-3 sediments was determined using an ITRAX X-ray Fluorescence Scanner (Cox Analytical Instruments). XRF scanning was conducted at 1-cm resolution with 60-s scans using a molybdenum X-ray source set to 30 kV and 15 mA.

Received 11 June 2010; accepted 12 January 2011.

- Woodhouse, C. A. *et al.* A 1,200-year perspective of 21st century drought in southwestern North America. *Proc. Natl Acad. Sci. USA* **107**, 21283–21288 (2010).
- McCabe, G. J., Palecki, M. A. & Betancourt, J. L. Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States. *Proc. Natl Acad. Sci. USA* **101**, 4136–4141 (2004).
- Cook, E. R., Seager, R., Cane, M. A. & Stahle, D. W. North American drought: reconstructions, causes and consequences. *Earth Sci. Rev.* **81**, 93–134 (2007).
- Seager, R. *et al.* Model projections of an imminent transition to a more arid climate in southwestern North America. *Science* **316**, 1181–1184 (2007).
- Weijers, J. W. H., Schouten, S., van den Donker, J. C., Hopmans, E. C. & Sinninghe Damsté, J. S. Environmental controls on bacterial tetraether membrane lipid distribution in soils. *Geochim. Cosmochim. Acta* **71**, 703–713 (2007).
- Berger, A. & Loutre, M. F. Insolation values for the climate of the last 10 million years. *Quat. Sci. Rev.* **10**, 297–317 (1991).
- Loutre, M. F. & Berger, A. Marine Isotope Stage 11 as an analog for the present interglacial. *Global Planet. Change* **36**, 209–217 (2003).
- Weijers, J. W. H., Schefuß, E., Schouten, S. & Sinninghe Damsté, J. S. Coupled thermal and hydrological evolution of tropical Africa over the last deglaciation. *Science* **315**, 1701–1704 (2007).
- McManus, J. F., Oppo, D. W. & Cullen, J. L. A 0.5-million-year record of millennial-scale climate variability in the North Atlantic. *Science* **283**, 971–975 (1999).
- Kutzbach, J. E. *et al.* Climate and biome simulation for the past 21000 years. *Quat. Sci. Rev.* **17**, 473–509 (2000).
- Prokopenko, A. A., Williams, D. F., Karabanov, E. B. & Khursevich, G. K. Response of Lake Baikal ecosystem to climate forcing and pCO₂ change over the Last Glacial/Interglacial transition. *Earth Planet. Sci. Lett.* **172**, 239–253 (1999).
- Prokopenko, A. A. *et al.* Muted climate variations in continental Siberia during the mid-Pleistocene epoch. *Nature* **418**, 65–68 (2002).
- de Vernal, A. & Hillaire-Marcel, C. Natural variability of Greenland climate, vegetation, and ice volume during the past million years. *Science* **320**, 1622–1625 (2008).
- Guo, Z. T., Berger, A., Yin, Q. Z. & Qin, L. Strong asymmetry of hemispheric climates during MIS-13 inferred from correlating China loess and Antarctic ice records. *Clim. Past* **5**, 21–31 (2009).
- Rosignol-Strick, M., Paterne, M., Bassinot, F. C., Emeis, K. C. & De Lange, G. J. An unusual mid-Pleistocene monsoon period over Africa and Asia. *Nature* **392**, 269–272 (1998).
- Jouzel, J. *et al.* Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science* **317**, 793–796 (2007).
- Loulergue, L. *et al.* Orbital and millennial-scale features of atmospheric CH₄ over the past 800,000 years. *Nature* **453**, 383–386 (2008).
- Johnson, D. W., Hanson, P. J., Todd, D. E., Susfalk, R. B. & Trettin, C. F. Precipitation change and soil leaching: field results and simulations from Walker Branch Watershed, Tennessee. *Wat. Air Soil Pollut.* **105**, 251–262 (1998).
- Meyers, P. A. Applications of organic geochemistry to paleolimnological reconstructions: a summary of examples from the Laurentian Great Lakes. *Org. Geochem.* **34**, 261–289 (2003).
- Ehleringer, J. R., Cerling, T. E. & Helliker, B. R. C₄ photosynthesis, atmospheric CO₂ and climate. *Oecologia* **112**, 285–299 (1997).
- Huang, Y. *et al.* Climate change as the dominant control on glacial-interglacial variations in C₃ and C₄ plant abundance. *Science* **293**, 1647–1651 (2001).
- Douglas, M. W., Maddox, R. A., Howard, K. & Reyes, S. The Mexican monsoon. *J. Clim.* **6**, 1665–1677 (1993).
- Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D. & Bacmeister, J. T. On the cause of the 1930s dust bowl. *Science* **303**, 1855–1859 (2004).
- Anderson, R. S. *et al.* Development of the mixed conifer forest in northern New Mexico and its relationship to Holocene environmental change. *Quat. Res.* **69**, 263–275 (2008).
- Jiménez-Moreno, G., Fawcett, P. J. & Anderson, R. S. Millennial- and centennial-scale vegetation and climate changes during the Late Pleistocene and Holocene from northern New Mexico (USA). *Quat. Sci. Rev.* **27**, 1448–1452 (2008).
- Poore, R. Z., Pavich, M. J. & Grissino-Mayer, H. D. Record of the North American southwest monsoon from Gulf of Mexico sediment cores. *Geology* **33**, 209–212 (2005).
- Enzel, Y., Cayan, D. R., Anderson, R. Y. & Wells, S. G. Atmospheric circulation during Holocene lake stands in the Mojave Desert: evidence of regional climate change. *Nature* **341**, 44–47 (1989).
- Namias, J. Some meteorological aspects of drought with special reference to the summers of 1952–54 over the United States. *Mon. Weath. Rev.* **83**, 199–205 (1955).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Mets for analytical support, W. McIntosh for the Ar–Ar age determination, T. Wawrzyniec and A. Ellwein for drilling help, and the Valles Caldera Trust for permission to drill in the Valle Grande. Core assistance was provided by LRC/LacCore. This work was supported by the NSF Paleoclimate and P2C2 programs, IGPP LANL and the USGS Western Mountain Initiative. Support from the Gladden Fellowship is acknowledged. This work forms contribution 2399-JW at the Centre for Water Research, The University of Western Australia and contribution 131 at the Laboratory of Paleoclimatology, Northern Arizona University.

Author Contributions Writing and interpretation was done by P.J.F. with significant contributions from J.P.W., R.S.A., J.M.H. and E.T.B. MBT/CBT analyses were conducted by J.P.W., M.A.B., J.S.S.D., S.S., Y.H. and J.T. Organic carbon/nitrogen analyses were conducted by P.J.F., J.M.H., L.M.C.-D., J.F. and V.A. XRF core scanning analyses were conducted by E.T.B. Pollen analyses and palaeovegetation analyses were conducted by R.S.A., S.J.S. and C.D.A., and F.G., G.W. and P.J.F. conducted core sediment and stratigraphic analyses. L.D.-H. and J.W.G. investigated palaeomagnetic and rock magnetic core properties. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.J.F. (fawcett@unm.edu).

Seismic tremors and magma wagging during explosive volcanism

A. Mark Jellinek¹ & David Bercovici²

Volcanic tremor is a ubiquitous feature of explosive eruptions. This oscillation persists for minutes to weeks and is characterized by a remarkably narrow band of frequencies from about 0.5 Hz to 7 Hz (refs 1–4). Before major eruptions, tremor can occur in concert with increased gas flux and related ground deformation^{5–7}. Volcanic tremor is thus of particular value for eruption forecasting^{6,8}. Most models for volcanic tremor rely on specific properties of the geometry, structure and constitution of volcanic conduits as well as the gas content of the erupting magma. Because neither the initial structure nor the evolution of the magma-conduit system will be the same from one volcano to the next, it is surprising that tremor characteristics are so consistent among different volcanoes. Indeed, this universality of tremor properties remains a major enigma. Here we employ the contemporary view that silicic magma rises in the conduit as a columnar plug surrounded by a highly vesicular annulus of sheared bubbles^{9,10}. We demonstrate that, for most geologically relevant conditions, the magma column will oscillate or ‘wag’ against the restoring ‘gas-spring’ force of the annulus at observed tremor frequencies. In contrast to previous models, the magma-wagging oscillation is relatively insensitive to the conduit structure and geometry, which explains the narrow band of tremor frequencies observed around the world. Moreover, the model predicts that as an eruption proceeds there will be an upward drift in both the maximum frequency and the total signal frequency bandwidth, the nature of which depends on the explosivity of the eruption, as is often observed.

The main characteristics of volcanic tremor depend strongly on whether a volcano is erupting explosively and on the intensity of the event (Fig. 1). Long before an eruption, tremor is ‘narrow-band’ (about 0.5–2 Hz and sometimes monochromatic or harmonic). On transitions to active volcanism, however, the maximum frequency can climb to 5–7 Hz (refs 1–4) (Fig. 1c and d). Moreover, whereas tremor related to low-intensity volcanism remains narrow-band, tremor associated with intermittent or protracted explosive behaviour is ‘broadband’, characterized by power distributed over the full ~0.5–7 Hz tremor bandwidth.

A successful model for volcanic tremor must ultimately explain its excitation, longevity and restricted frequency bandwidth, as well as its evolution in spectral character over changes in eruption style (Fig. 1). Such an endeavour is difficult because the signal can include contributions from the source region (‘source effects’) and the magma-filled conduit system through which seismic waves propagate (‘path effects’)^{1,11–13}. Previous modelling efforts suggest that tremor can be driven and modulated by processes related to time-dependent mechanical coupling of ascending magma to the conduit walls^{14,15}, unsteady stirring in a gas-rich magma^{8,16,17}, and the interaction of ascending magma with irregularities in the conduit diameter and wall structure, such as constrictions¹⁸ and cracks^{8,19}. The initial signal bandwidth and its upward drift with volcanism might be related to time-varying path effects, such as the gas content, bubble size distribution and bubble nucleation depth in the magma, as well as the height and average

diameter of the conduit^{1,11–13,17}. Alternatively, time-dependent source effects—including the location, shape and structure of cracks of various class within the conduit walls, which act as ‘resonators’^{1,20}, and the depth of seismicity related to brittle failure in the magma itself^{14,21,22}—may influence the signal at the surface. A challenge for these models is that the structure and constitution of magma-conduit systems need not be similar from one volcano to the next before or during an eruption. Indeed, differences among the styles of conduit erosion^{4,23} and magma degassing⁹ would be predicted to produce large variations in the tremor signal bandwidth that are not observed (Fig. 1). In addition, it is not clear why the signal climbs in frequency but remains almost harmonic for low-energy events whereas it becomes broadband on transitions to Vulcanian or Plinian eruptive styles (Fig. 1).

Although the magma-conduit system may vary among volcanoes, for silica-rich magmas characteristic of explosive systems, experiments and calculation show that the magma—a mixture of very viscous melt, crystals and bubbles—rises as a stiff columnar plug that is highly sheared by viscous drag at the conduit walls¹⁰. In these systems bubbles and crystals do not move relative to the melt and impart a strain-rate-dependent rheology to the magma mixture¹⁰. In particular, depending on the gas content of the magma, shear strains are localized near the walls as a result of the preferential deformation of essentially inviscid bubbles, which causes the magma to flow as a plug. The depth H over which a plug flow is appropriate depends on the crystal content and the concentration and solubility of volcanic gas (primarily H₂O) and is of the order of a kilometre^{10,24}. A key consequence of this plug flow is that the zone of sheared bubbles at the conduit wall develops into a thin, highly vesicular, permeable and compressible ‘annulus’ enveloping a ‘stiff’ central magma plug (Fig. 2)^{9,10,25,26}. Textural analyses of pumice from effusive and explosive eruptions^{27,28} show that the sheared and elongated bubbles within the annulus form a partially connected tube-like network with porosity ϕ_0 of 30–90% (Fig. 2). Thus, the annulus is also permeable^{27,28}, so that gas can ascend relative to the magma column near the conduit wall^{9,25}. Indeed, observations of rings of gas rising around the edges of active volcanic vents²⁶ (Fig. 2) suggest the presence of such a permeable annulus through which gas is transported.

Given the magma column–annulus structure, volcanic tremor would be a plausible consequence of the magma column rattling or wagging inside the bubble-rich annulus. In particular, lateral displacements of the column from a resting position in the centre of the conduit will compress or dilate the annulus, depending on which way the column is displaced. The annulus is essentially a semi-permeable springy foam that acts to restore the magma column to its resting position; however, the column’s inertia causes it to overshoot that position, in which case an oscillation emerges and the column ‘wags’ back and forth. The corresponding gas-pressure fluctuations in the annulus are transmitted to the walls, giving rise to the tremor observed with seismometers and infrasound sensors at the surface.

A theoretical model describing a magma column oscillating within a compressible annulus is presented in the Supplementary Information. The essential physics of magma wagging is, however, captured by a

¹Department of Earth and Ocean Sciences, The University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada. ²Department of Geology & Geophysics, Yale University, New Haven, Connecticut 06511, USA.

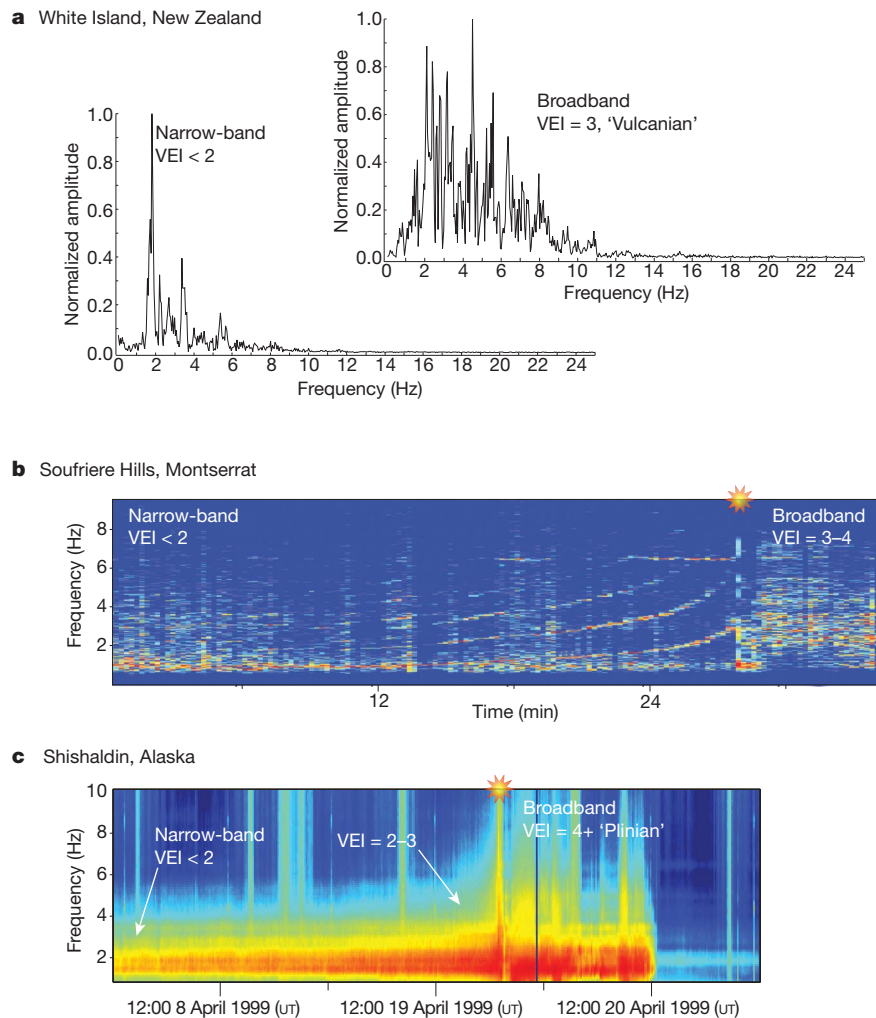


Figure 1 | Variation of tremor signal with eruption intensity. Tremor properties vary depending on the intensity of explosive eruptions, which is classified in terms of the Volcanic Explosivity Index (VEI). High-intensity so-called Vulcanian and Plinian events (VEI > 3) are characterized by catastrophic and potentially protracted fragmentation of the magma indicated by tall eruption columns. Low-intensity (VEI < 3) events are not associated with widespread magma fragmentation. **a**, Amplitude spectra for tremor from a 1998 eruption of White Island, New Zealand, showing a widening of the tremor signal bandwidth following a transition from 'normal' background volcanic unrest to more energetic Vulcanian behaviour (modified from refs 2 and 29) after an explosion. The noise floor above which frequencies can be visually distinguished increases from a pre-eruptive value close to the baseline to

perhaps 20% of the signal during eruption, consistent with observations at other volcanoes²³. **b**, **c**, Time series of amplitude spectra from eruptions at Montserrat (**b**, modified from ref. 1) and Shishaldin (**c**, modified from ref. 30) showing evolutions in tremor characteristics during transitions from background volcanic unrest to eruptions of various intensities. As with the White Island example, tremor begins as a harmonic or nearly monochromatic ('narrow-band') signal restricted to a signal bandwidth of 0.5–2 Hz. As the eruptions progress, both the maximum frequency and the total signal bandwidth increase to become 'broadband' over a bandwidth of 0.5–7 Hz. Whereas tremor during low-intensity eruptions can remain narrow-band, it is broadband for high intensity (VEI > 3) explosive volcanism.

simplified version of the model in which the annulus contains a compressible but impermeable gas–magma foam of disconnected bubbles that rises with the magma column (see Methods). The behaviour of a column of very viscous magma surrounded by a foam-spring of fixed mass is similar to an unforced, damped harmonic oscillator. Indeed, the natural angular frequency of the wagging magma column is the primary model prediction for volcanic tremor and is given by:

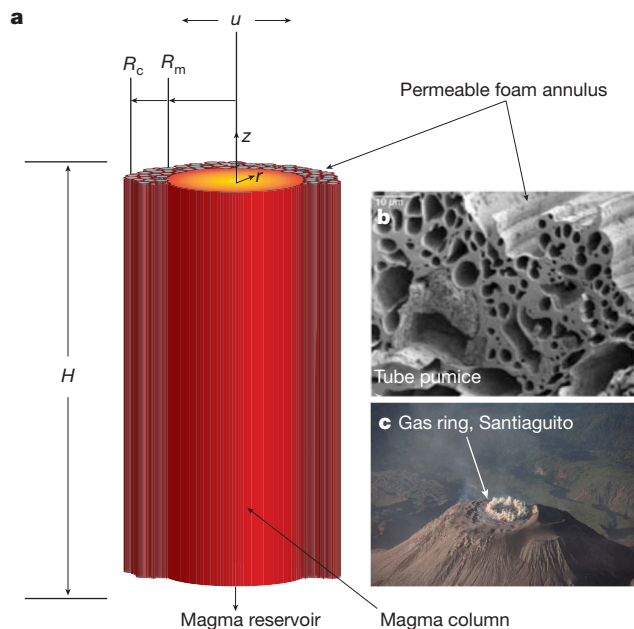
$$\omega_o = \sqrt{\frac{2\rho_o C^2}{\phi_o \rho_m (R_c^2 - R_m^2)}} \quad (1)$$

where the ordinary frequency is $\omega_o/(2\pi)$, C is the speed of sound in gas (appropriate for water vapour at magmatic temperatures), ρ_o and ϕ_o are the density and volume fraction of the undisturbed gas in the annulus, ρ_m is the magma density, and R_c and R_m are the conduit and magma column radii, respectively (see Methods and Fig. 2).

Although there are several properties embedded in the frequency relation (1), most are well constrained and the main uncertainty and

likely variation among volcanoes is geometrical, that is, in the radii R_c and R_m . However, because $\omega_o \propto (R_c^2 - R_m^2)^{-1/2}$, the effect of uncertainty in the system geometry (or its evolution during an eruption) is relatively small. For a reasonably complete set of possible conduit properties, the predicted wagging frequency $\omega_o/(2\pi)$ varies between approximately 0.1 and 5 Hz (Fig. 3), which is similar to the observed bandwidth of ~0.5–7 Hz for volcanic tremor. A more complete analysis including the vertical flow of gas in the annulus does not alter this key result (see Supplementary Information). For plausible eruption conditions, magma wagging will excite tremor at the observed frequencies.

The longevity of tremor is another remarkable feature of this phenomenon. In the model, the persistence of the wagging oscillation depends on damping related to the viscous bending and shearing of the magma column (see Methods and Supplementary Information). The amount of shearing depends on the ratio of the wagging velocity to the column height and so oscillations persist longer for taller columns. Indeed, for wagging oscillations to occur, the column must exceed a critical height:



$$H_c = \frac{\pi}{2} \sqrt{\frac{\mu_m}{2\rho_m \omega_0}} \quad (2)$$

where μ_m is the magma viscosity (see Methods). For typical viscosities for explosively erupting silicic magmas $10^5 \text{ Pa s} \leq \mu_m \leq 10^7 \text{ Pa s}$ (ref. 10) and wagging frequencies $0.5 \text{ Hz} \leq \omega_0/(2\pi) \leq 7 \text{ Hz}$ (Fig. 3), and taking $\rho_m = 2,500 \text{ kg m}^{-3}$, the column must be over 40 m high for wagging to occur, which is far below the typical 1-km height H for the magma column. The magma wagging oscillation, however, is invariably damped and must be driven by some excitation or forcing if it is to persist for the maximum observed tremor duration. In most systems a requisite impulsive excitation can occur by various episodic mechanisms¹ including fluctuations in the gas flux, explosions (Fig. 1, red

stars), rockfalls and (non-tremor-related) volcano-tectonic seismic events^{2,3,21,22}, as well as by processes related to unsteady magma flow and fragmentation^{8,15,16,21}. This model can also be used to explain the increase in tremor frequency with the onset of volcanism as well as the broadband tremor spectrum that characterizes transitions to explosive Plinian-style behaviour (Fig. 1). As noted above, eruptions typically follow a period of volcano pressurization and enhanced gas flux, indicating a pre-eruptive build-up of gas bubbles in the magma. Near the conduit walls, a steady increase in the bubble concentration will enhance the localization of shear deformation¹⁰, causing the annulus to narrow progressively, the gas pressure to rise and the wagging frequency (and associated harmonics) to climb, in turn (see equation (1) and Fig. 3), as is sometimes observed (Fig. 1b and c). At a critical annular shear strain rate, which is proportional to the inverse of the declining annulus thickness, the sheared bubbles will rupture catastrophically or fragment^{10,22}, and the annulus will evolve rapidly from a foam composed of tube-like bubbles to a mixture of fractured tubes and shattered glass fragments. Fragmentation would therefore coincide with the minimum thickness

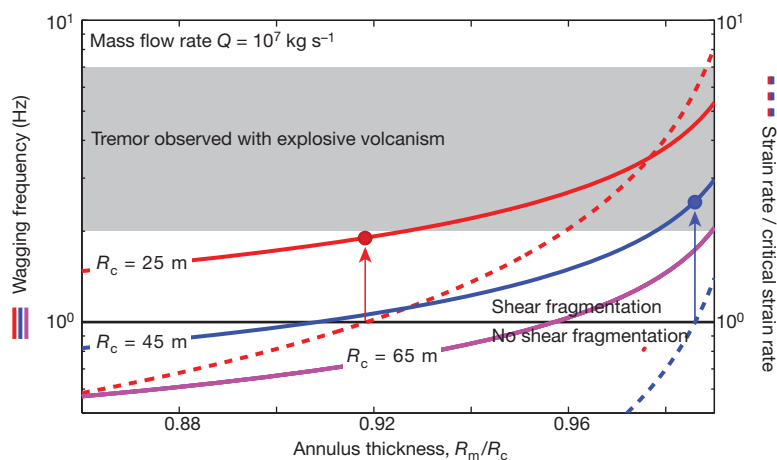


Figure 3 | Wagging frequency and shear strain rate in the annulus versus annulus thickness for several conduit radii. Solid lines indicate wagging frequency $\omega_0/(2\pi)$ and dashed lines indicate shear strain rates. The grey shaded region indicates the range of frequencies associated with explosive volcanism. Lower wagging frequencies correspond to wider columns because of their greater inertia. Shear strain rates are normalized to the critical value for shear brecciation or fragmentation. The black line corresponding to a normalized strain rate of unity is the fragmentation threshold. An increase in the gas content in the magma before or during an eruption will increase the bubble concentration near the conduit walls, thereby enhancing shear strain localization and causing the annulus to narrow in turn. Whereas the wagging frequency increases as $(R_c^2 - R_m^2)^{-1/2}$ (see equation (1)), the strain rate increases more rapidly, as $(R_c - R_m)^{-1}$, towards a critical value at which bubble

walls rupture and the magma fragments. Arrows show the conditions at which shear fragmentation is expected and the corresponding wagging frequency. For the plausible range in conduit radii shown, the wagging frequency at fragmentation is comparable to that which is observed (Fig. 1). Following fragmentation, the gas spring force will be spatially variable, leading potentially to an increase in the bandwidth of fundamental modes, depending on the wagging direction (see text). The strain rate is $(Q/(\rho_m \pi R_c^2))/(R_c - R_m)$ and the critical strain rate for shear fragmentation is $(\gamma G/\mu_m)$, where Q is a typical mass flow rate¹⁰, $G = 10^{10} \text{ Pa}$ is the shear modulus of the melt and $\gamma = 10^{-3}$ is a geometric constant. For bubbly magmas it is expected that $\gamma < 10^{-2}$ (ref. 10) and we have adjusted the value such that a 3 Hz tremor is recovered at fragmentation, consistent with Fig. 1c.

of the annulus, and thus the maximum wagging frequency. Indeed, for plausible conduit geometries the wagging frequencies associated with fragmentation are consistent with the maximum tremor frequencies that have been observed to follow transitions to Vulcanian and Plinian volcanism (Fig. 1b and c; Fig. 3). Moreover, fragmentation will introduce strong spatial heterogeneity into the structure of the annulus and into the behaviour of the gas spring. Consequently the natural wagging oscillation for the magma column will plausibly be characterized by many frequencies, leading to the relatively broadband tremor signal observed (Fig. 1).

Volcanic tremor is a remarkably consistent feature of explosive volcanism observed around the world. Although tremor is critical for predicting and characterizing hazardous volcanic eruptions³, the physical understanding for its origin has been elusive. We have shown, however, that the full bandwidth of observed tremor frequencies, as well as their evolution during eruption, can be explained by the dynamics of a simple magma–‘wagging’ oscillation. Therefore, tremor may be a natural consequence of the well-established plug-like structure of the conduit flow anticipated for most of these systems. In contrast to previous tremor theories, our model predicts the observed narrow frequency band for a wide range of explosive volcanic settings and conduit geometries. The proposed magma-wagging model therefore provides a fundamental mechanism for tremor that is generic to nearly all volcanically explosive systems.

METHODS SUMMARY

The theoretical results for the wagging magma column are based on two approaches: first an idealized fluid-dynamical model of a magma column oscillating within an annulus of impermeable magma foam; and second, a model in which the annulus is filled with a permeable matrix of sheared, elongated and interconnected bubbles or tubes through which the gas migrates relative to the magma. In either case, the oscillation of the magma column is governed by the inertia of the column, the restoring gas-spring force of the annulus, and the viscous resistance to shearing or bending of the column. The gas-spring force is determined from variations in gas volume and pressure in the annulus caused by its lateral distortions during oscillations of the magma column (that is, contraction of the annulus ahead of the column's displacement, and dilation behind it). The bending resistance is determined through the viscous stress associated with horizontal shearing in the magma column. The sum of these forces governs the lateral acceleration of the magma column and leads to a governing equation for the oscillations that can be derived from Newton's second law. Thus, both the foam and permeable annulus models describe the coupling of column displacement with flow and compression of the gas in the annulus. The fundamental oscillation frequency and damping rate are the same for both models, as described by the linearized governing equations. Therefore, because the simpler foam model captures much of the essential physics of the system, it is the basis of discussion here and is developed in Methods. A comprehensive development of the theory, including the permeable annulus model, is presented in Supplementary Information.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 August 2010; accepted 7 January 2011.

1. Neuberg, J. Characteristics and causes of shallow seismicity in andesite volcanoes. *Phil. Trans. Math. Phys. Engin. Sci.* **358**, 1533–1546 (2000).
2. Konstantinou, K. & Schlindwein, V. Nature, wavefield properties and source mechanism of volcanic tremor: a review. *J. Volcanol. Geotherm. Res.* **119**, 161–187 (2003).
3. McNutt, S. Volcanic seismology. *Annu. Rev. Earth Planet. Sci.* **33**, 461–491 (2005).
4. McNutt, S. & Nishimura, T. Volcanic tremor during eruptions: temporal characteristics, scaling and constraints on conduit size and processes. *J. Volcanol. Geotherm. Res.* **178**, 10–18 (2008).
5. Dzurisin, D. A comprehensive approach to monitoring volcano deformation as a window on the eruption cycle. *Rev. Geophys.* **41**, doi:10.1029/2001RG000107 (2003).
6. Sparks, R. Forecasting volcanic eruptions. *Earth Planet. Sci. Lett.* **210**, 1–15 (2003).

7. Lu, Z., Dzurisin, D., Biggs, J., Wicks, C. & McNutt, S. Ground surface deformation patterns, magma supply, and magma storage at Okmok volcano, Alaska, from InSAR analysis: 1. Interruption deformation, 1997–2008. *J. Geophys. Res.* **115**, B00B02, doi:10.1029/2009JB006969 (2010).
8. Chouet, B. Long-period volcano seismicity: its source and use in eruption forecasting. *Nature* **380**, 309–316 (1996).
9. Sparks, R. Dynamics of magma degassing. *Geol. Soc. Lond. Spec. Publ.* **213**, 5–22 (2003).
10. Gonnermann, H. & Manga, M. The fluid mechanics inside a volcano. *Annu. Rev. Fluid Mech.* **39**, 321–356 (2007).
11. Chouet, B. *et al.* Source and path effects in the wave fields of tremor and explosions at Stromboli Volcano, Italy. *J. Geophys. Res.* **102**, 15129–15150 (1997).
12. Sherburn, S., Bryan, C., Hurst, A., Latter, J. & Scott, B. Seismicity of Ruapehu volcano, New Zealand, 1971–1996: a review. *J. Volcanol. Geotherm. Res.* **88**, 255–278 (1999).
13. Hagerty, M., Schwartz, S., Garcés, M. & Protti, M. Analysis of seismic and acoustic observations at Arenal Volcano, Costa Rica, 1995–1997. *J. Volcanol. Geotherm. Res.* **101**, 27–65 (2000).
14. Goto, A. A new model for volcanic earthquake at Unzen Volcano: melt rupture model. *Geophys. Res. Lett.* **26**, 2541–2544 (1999).
15. Denlinger, R. & Hoblitt, R. Cyclic eruptive behavior of silicic volcanoes. *Geology* **27**, 459–462 (1999).
16. Hellweg, M. Physical models for the source of Lascar's harmonic tremor. *J. Volcanol. Geotherm. Res.* **101**, 183–198 (2000).
17. Garcés, M., A., Hagerty, M. T. & Schwartz, S. Y. Magma acoustics and time-varying melt properties at Arenal Volcano, Costa Rica. *Geophys. Res. Lett.* **25**, 2293–2296 (1998).
18. Julian, B. Volcanic tremor: nonlinear excitation by fluid flow. *J. Geophys. Res.* **99**, 11859–11877 (1994).
19. Rust, A., Balmforth, N. & Mandre, S. The feasibility of generating low-frequency volcano seismicity by flow through a deformable channel. *Geol. Soc. Lond. Spec. Publ.* **307**, 45–56 (2008).
20. Hurst, A. & Sherburn, S. Volcanic tremor at Ruapehu: characteristics and implications for the resonant source. *NZ J. Geol. Geophys.* **36**, 475–485 (1993).
21. Neuberg, J. *et al.* The trigger mechanism of low-frequency earthquakes on Montserrat. *J. Volcanol. Geotherm. Res.* **153**, 37–50 (2006).
22. Tuffen, H., Smith, R. & Sammonds, P. Evidence for seismogenic fracture of silicic magma. *Nature* **453**, 511–514 (2008).
23. Bryan, C. & Sherburn, S. Seismicity associated with the 1995–1996 eruptions of Ruapehu volcano, New Zealand: narrative and insights into physical processes. *Bull. Volcanol.* **65**, 30–42 (2003).
24. Collier, L., Neuberg, J., Lensky, N., Lyakhovsky, V. & Navon, O. Attenuation in gas-charged magma. *J. Volcanol. Geotherm. Res.* **153**, 21–36 (2006).
25. Eichelberger, J. C., Carrigan, C. R., Westrich, H. R. & Price, R. H. Non-explosive silicic volcanism. *Nature* **323**, 598–602 (1986).
26. Bluth, G. & Rose, W. Observations of eruptive activity at Santiaguito volcano, Guatemala. *J. Volcanol. Geotherm. Res.* **136**, 297–302 (2004).
27. Klug, C. & Cashman, K. Permeability development in vesiculating magmas: implications for fragmentation. *Bull. Volcanol.* **58**, 87–100 (1996).
28. Wright, H., Cashman, K., Gottesfeld, E. & Roberts, J. Permeability of anisotropic tube pumice: model calculations and measurements. *Earth Planet. Sci. Lett.* **280**, 93–104 (2009).
29. Sherburn, S., Scott, B., Nishi, Y. & Sugihara, M. Seismicity at White Island volcano, New Zealand: a revised classification and inferences about source mechanism. *J. Volcanol. Geotherm. Res.* **83**, 287–312 (1998).
30. Thompson, G., McNutt, S. & Tytgat, G. Three distinct regimes of volcanic tremor associated with the eruption of Shishaldin Volcano, Alaska 1999. *Bull. Volcanol.* **64**, 535–547 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements A.M.J. acknowledges support from the Canadian Institute for Advanced Research and NSERC. D.B. acknowledges support from the National Science Foundation. This manuscript has benefited from discussions and comments from S. McNutt and C. Michaut. We thank S. McNutt, J. Neuberg, M. Hagerty, K. I. Konstantinou and M. Ibs-von Seht for data and preliminary figures that were incorporated into various versions of Fig. 1.

Author Contributions A.M.J. and D.B. conceived the physical model for magma wagging together. D.B. developed the mathematical model. A.M.J. collected and analysed the seismic and acoustic data, developed the model for fragmentation applied in Fig. 3 and was the lead author for the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.M.J. (mjellinek@eos.ubc.ca).

METHODS

Foam annulus model. In the simplified model, the magma plug or column wags inside a compressible annular jacket of impermeable foam with a high volume fraction of disconnected gas bubbles. The foam moves with the ascending magma column, to which we affix the reference frame.

The inner magma column is assumed to be cylindrical with a radius R_m , and when at rest is centred inside a cylindrical conduit of radius R_c (see Fig. 2). The horizontal displacement of the column at any height z is $u(z)$. The net pressure force exerted by the foam annulus on the magma column is governed by how much the foam is squeezed or dilated by the magma wagging. In a polar coordinate system centred on the axis of the conduit, the magma column's cylindrical surface is given by $(r_m \cos \theta - u)^2 + (r_m \sin \theta)^2 = R_m^2$, where r_m is the distance from the conduit axis to the magma column's surface, and θ is the angle relative to the direction of displacement. Assuming small displacements in which $u \ll R_m$, we obtain:

$$r_m(\theta) = R_m + u \cos \theta \quad (3)$$

The volume of a segment of foam in the annulus from θ to $\theta + d\theta$ and of vertical height dz is:

$$V = dz d\theta \int_{r_m(\theta)}^{R_c} r dr = \frac{1}{2} dz d\theta (R_c^2 - R_m^2 - 2u R_m \cos \theta) \quad (4)$$

The undisturbed volume of this segment is $V_0 = \frac{1}{2} dz d\theta (R_c^2 - R_m^2)$, and hence the general volume can be cast as:

$$V = V_0 \left(1 - \frac{u}{U} \cos \theta \right) \quad (5)$$

where $U = \frac{R_c^2 - R_m^2}{2R_m}$. The gas volume V_g is obtained from the fact that the volume of magma in the foam annulus is conserved, hence $V - V_g = V_0 - V_{g0}$, where V_{g0} is the unperturbed gas volume. This leads to:

$$V_g = V_0 \left(\phi_0 - \frac{u}{U} \cos \theta \right) \quad (6)$$

where $\phi_0 = V_{g0}/V_0$ is the unperturbed gas volume fraction. The foam pressure is related to the gas volume by the ideal gas law:

$$P = \frac{(M_0/m_g) \mathcal{R} T}{V_g} \approx \rho_0 C^2 \left(1 + \frac{u}{\phi_0 U} \cos \theta \right) \quad (7)$$

where M_0 is the mass of gas in the foam segment, which is conserved, and thus $\rho_0 = M_0/(\phi_0 V_0)$ is the undisturbed gas density. \mathcal{R} is the gas constant, T is temperature, m_g is the molar mass of the gas, and $C = \sqrt{\mathcal{R} T / m_g}$ is the speed of sound in gas. The traction (force/area) of the gas pressure pushing on the magma column in the direction of displacement is:

$$-P \hat{\mathbf{x}} \cdot \hat{\mathbf{n}} \approx -P \left(\cos \theta - \frac{u}{R_m} \sin^2 \theta \right) \quad (8)$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{n}}$ are the unit vectors in the direction of displacement, and normal to the magma column's surface, respectively. Integrating this traction completely

around the annulus (that is, for $0 \leq \theta \leq 2\pi$) gives the net pressure force on a vertical segment of the magma column of thickness dz :

$$F_p = -dz \int_0^{2\pi} P(\theta) \hat{\mathbf{x}} \cdot \hat{\mathbf{n}} r_m(\theta) d\theta = -dz \rho_0 C^2 \frac{\pi R_m}{\phi_0 U} u \quad (9)$$

The resistance force due to viscous bending of the column is given by the difference in shear tractions between the top and bottom of a column segment of height dz , which is:

$$F_b = [\tau_{xz}]_z^{z+dz} \pi R_m^2 \quad (10)$$

where τ_{xz} is the shear stress resisting the magma column bending to the left or right. At any given point in the column we can approximate (for small displacement) the stress as:

$$\tau_{xz} = \mu_m \frac{\partial^2 u}{\partial z \partial t} \quad (11)$$

where μ_m is the magma's dynamic viscosity.

Newton's second law for the segment of the magma column relates the column segment's mass $\rho_m \pi R_m^2 dz$, acceleration $\partial^2 u / \partial t^2$ and net force $F_p + F_b$ via:

$$\rho_m \pi R_m^2 dz \frac{\partial^2 u}{\partial t^2} = -dz \rho_0 C^2 \frac{\pi R_m}{\phi_0 U} u + [\tau_{xz}]_z^{z+dz} \pi R_m^2 \quad (12)$$

which, in the limit of small dz , and using equation (11), eventually becomes:

$$\frac{\partial^2 u}{\partial t^2} = -\omega_0^2 u + \frac{\mu_m}{\rho_m} \frac{\partial^3 u}{\partial z^2 \partial t} \quad (13)$$

where

$$\omega_0^2 = \frac{\rho_0 C^2}{\phi_0 \rho_m R_m U} = \frac{2\rho_0 C^2}{\phi_0 \rho_m (R_c^2 - R_m^2)} \quad (14)$$

leads to the relation for the natural angular frequency of magma-wagging shown in equation (1).

Assuming $u \approx e^{i(\omega_0/C)kz + st}$, where $k = 2\pi C/(\omega_0 \lambda)$ is the dimensionless wave-number associated with a magma-wagging wavelength λ , and s is the oscillation growth rate, one arrives at the dispersion relation:

$$s/\omega_0 = -\eta k^2 \pm \sqrt{\eta^2 k^4 - 1} \quad (15)$$

where $\eta = \mu_m \omega_0 / (2\rho_m C^2)$ expresses the balance between viscous bending and foam spring forces. For the special case $\eta k^2 \ll 1$ we obtain the imaginary growth rate $s \approx i\omega_0$, which correspond to undamped wagging. Oscillations will only occur for $\eta k^2 < 1$ or equivalently for $\lambda > 2\pi C \sqrt{\eta} / \omega_0$. Assuming a fundamental mode of wagging with the displacement maximized at the top surface (recognizing that higher modes will be damped faster), a full wavelength for the perturbation will be $\lambda = 4H$; thus H must exceed the critical value given by equation (2) for oscillations to occur. The longest dimensional damping time—that is, $(2H/\pi)^2 (2\rho_m/\mu_m)$, which is the inverse of the decay rate $-\text{Re}(s)$ from equation (15), assuming $\eta k^2 < 1$, and with $\lambda = 4H$ —for the other properties listed above and $H = 1$ km is about 6 hours.

An armoured Cambrian lobopodian from China with arthropod-like appendages

Jianni Liu^{1,2}, Michael Steiner², Jason A. Dunlop³, Helmut Keupp², Degan Shu^{1,4}, Qiang Ou⁴, Jian Han¹, Zhifei Zhang¹ & Xingliang Zhang¹

Cambrian fossil Lagerstätten preserving soft-bodied organisms have contributed much towards our understanding of metazoan origins^{1–3}. Lobopodians are a particularly interesting group that diversified and flourished in the Cambrian seas. Resembling ‘worms with legs’, they have long attracted much attention in that they may have given rise to both Onychophora (velvet worms)^{4–6} and Tardigrada (water bears)^{7,8}, as well as to arthropods in general^{9–12}. Here we describe *Diania cactiformis* gen. et sp. nov. as an ‘armoured’ lobopodian from the Chengjiang fossil Lagerstätte (Cambrian Stage 3), Yunnan, southwestern China. Although sharing features with other typical lobopodians, it is remarkable for possessing robust and probably sclerotized appendages, with what appear to be articulated elements. In terms of limb morphology it is therefore closer to the arthropod condition, to our knowledge, than any lobopodian recorded until now. Phylogenetic analysis recovers it in a derived position, close to Arthropoda; thus, it seems to belong to a grade of organization close to the point of becoming a true arthropod. Further, *D. cactiformis* could imply that arthropodization (sclerotization of the limbs) preceded arthroization (sclerotization of the body). Comparing our fossils with other lobopodian appendage morphologies—see *Kerygmachela*^{9,10}, *Jianshanopodia*¹³ and *Megadictyon*¹²—reinforces the hypothesis that the group as a whole is paraphyletic, with different taxa expressing different grades of arthropodization.

Lobopodia Snodgrass, 1938

Xenusia Dzik & Krumbiegel, 1989

Diania cactiformis gen. et sp. nov.

Etymology. *Dian* represents a linguistic abbreviation in Chinese of the Yunnan province, where the Chengjiang fossil Lagerstätte is located; *cactiformis* refers to its overall cactus-like shape, which resulted in its nickname ‘walking cactus’.

Holotype. ELI-WT006A,B (Fig. 1a–d), part and counterpart of the complete specimen, which was collected in 2006 from the Jianshan section, Haikou, Yunnan, southwestern China. ELI is an abbreviation of the Early Life Institute, Northwest University, Xi’an, China.

Locality. Yu’anshan (Heilinpu) Formation (*Wutingaspis*–*Eoredlichia* Zone).

Horizon. Cambrian Stage 3.

Diagnosis. Armoured lobopodian with ten pairs of appendages. Trunk region with nine segments, bearing rows of transverse annulations each with some tubercles. Each region possesses a pair of robust and sclerotized spiny appendages with primary articulation. Anterior is extended, probably forming a proboscis. Posterior region bears a protrusion.

Diania cactiformis is known from the part and counterpart of three complete specimens, and about 30 other incomplete specimens (see Methods). The complete specimens are about 6-cm long, with a proboscis-like structure bearing some wrinkles, unlike those of *Xenusia*¹⁴ or *Microdictyon*⁶. The proboscis of *Diania cactiformis* is slightly

expanded compared to the trunk and shows no sign of becoming thinner towards the end. No mouth can be observed (Figs 2a–e and 3). A small projection occurs at the posterior end of the body (Figs 2a–c and 3). The main body trunk is sub-circular in outline and composed of nine segments, each comprising five rows of sub-parallel transverse annulations alternating with five rings of tubercles with tiny spines (Fig. 1a, b); each with 10–12 tubercles. The tiny spines are not well-preserved, possibly as an artefact of taphonomy, and in most cases only the tubercles of these spines are observable (Figs 1a, b and 2b–c).

A pair of strongly armoured appendages protrudes from the ventrolateral region of each segment, each limb attached robustly at its base to the otherwise slender trunk. Indeed, there is usually a hard, disc-like structure observed in the middle of each trunk segment, with only one such disc per appendage pair (Figs 1a, b and 2a–c, g); different from, for example, the paired spine bases seen in *Hallucigenia*. Some filamentary lines occur within these disc-like structures. These lines are interpreted here as the remains of muscle tissue associated with the limb attachment area (Figs 1a, b and 2a–c, g). The attachment area is a little wider than the adjacent trunk segments. Most appendages are incompletely preserved, although in the holotype the seventh right appendage is very clear and is composed of some annulations at the base and sixteen strongly sclerotized segments (Fig. 1a, b). It is worth noting that adjacent segments show a ‘joint-like’ morphology (Figs 1a, b and 2f, h, i–l) and that each segment bears two spines on both lateral sides and two spines on the dorsal side (Figs 1a, b and 2h–k). Further, the appendages terminate in two spines (Fig. 2j). These spines, especially the two dorsal spines with their rounded bases (Figs 1a, b and 2i, k), robustly attach to the appendages. However, probably owing to dorsoventral flattening, the two dorsal spines are usually compressed and can appear shorter than the two lateral spines. These are thorn-like with a maximum height of 2.2 mm and a basal width of 1.1 mm. These spines seem rather stiff and conical in shape, as implied by their triangular profile with a cone angle of approximately 30° (Fig. 2h–k). There is a fold-like structure between the bottom of the triangle and the rim of the appendage (Figs 1a, b and 2j, l), which indicates that the lateral spines also have a solid attachment to the trunk. In particular, the angle between the appendages and trunk differs going from front to back. The angle between the appendages and the trunk in the holotype increases from ~5° (first pair of appendages) to ~10° (second pair) to ~20° (third pair) to ~40° (fourth pair), and then increases abruptly such that nearly all are at right angles (Figs 1a, b and 3). This implies that the appendages of the anterior and posterior body regions may have differed in their function. The anterior appendages were more ventrally orientated and thus more suitable for grasping/feeding from the substrate, whereas the posterior appendages were more laterally orientated and may have served more for locomotion. In specimen ELI-WD001A, a light black band is observed running from the base of the appendage towards the tip. We suggest that this was an extension of the body cavity or gut caeca (Fig. 2f), and that corresponding to each

¹Early Life Institute, State Key Laboratory of Continental Dynamics, Department of Geology, Northwest University, Xi’an 710069, China. ²Department of Earth Science, Freie Universität Berlin, D-12249 Berlin, Germany. ³Museum für Naturkunde, Leibniz Institute for Research on Evolution and Biodiversity at the Humboldt University Berlin, D-10115 Berlin, Germany. ⁴School of Earth Sciences and Resources, China University of Geosciences, Beijing 100083, China.

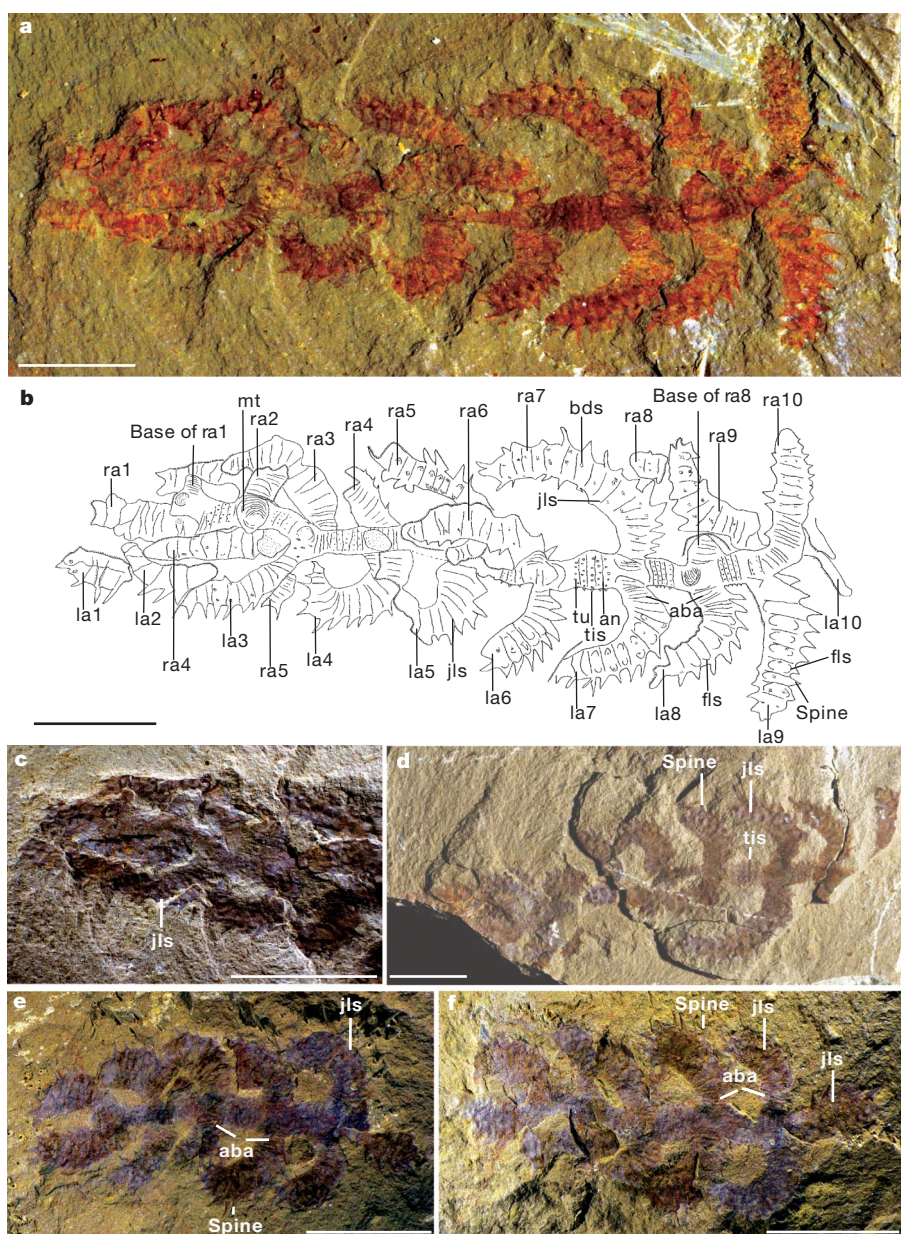


Figure 1 | The early Cambrian lobopodian *Diania cactiformis* gen. et sp. nov. from the Chengjiang fossil Lagerstätte, Kunming, Yunnan, China.

a–d, Holotype ELI-WD006A,B. **a**, Part of holotype photographed under ethanol, revealing almost all morphological features; note that the specimen is twisted towards the anterior end and the appendages are consequently slightly disordered, especially the fourth right appendage protruding under the trunk and appearing on the left side and the fifth right appendage protruding above the trunk and appearing again on the left side. **b**, Camera lucida drawing of **a**. **c**, Detail of the anterior part of **a**, which shows that the appendages are not

segment making up the appendage this cavity or gut caeca is divided into several small parts by joint-like structures (Fig. 2i–l), similar to that of Euarthropoda. In specimen ELI-WD0026A,B, a 2-mm-long and 1-mm-wide trunk extension is observed on the posterior-most body segment (Figs 2a–c and 3), similar to that of *Miraluolishania*¹¹ and *Hallucigenia fortis*⁶. We interpret this structure as a posterior protrusion as well.

Recent Onychophora (velvet worms) are frequently considered to be an outgroup of arthropod phylogeny, sharing with them numerous synapomorphies (as reviewed in ref. 15). Given that fossil lobopodians are almost certainly related in some way to velvet worms^{5,6,16}, they represent a particularly valuable group for understanding the transition from soft-bodied ecdysozoans to fully sclerotized arthropods.

preserved in the same layer; probably the specimens are to some extent three-dimensional. **d**, Counterpart of holotype. **e**, **f**, Specimen ELI-WT002A,B; although the specimens are incomplete, the appendages have clear annulations at the base and joint-like structures. **aba**, annulations of the base of appendages; **an**, annulations; **bds**, base of the dorsal spines; **dls**, disc-like structure; **fls**, fold-like structures; **jls**, joint-like structures; **la**, left appendage; **mt**, muscle tissue; **pr**, proboscis; **ra**, right appendage; **pp**, posterior protrusion; **tis**, tiny spine; **tu**, tubercle; **wr**, wrinkle. Scale bars, 10 mm.

Key innovations towards the Arthropoda condition (Euarthropoda in the sense of ref. 17) are eyes, sclerotized bodies and sclerotized, jointed limbs—whereas crown-group Arthropoda can be defined by, for example, a distinct head and biramous limbs with an endopod composed of relatively few articles.

How close Cambrian lobopodians are to the ground plan of the arthropod common ancestor remains a point of debate^{10,18}. Yet it is worth reiterating that in living Onychophora it is the claws at the tips of the legs that are fully sclerotized, whereas numerous fossil lobopodians preserve plates or spines ornamenting the trunk⁶. This begs the question whether the transition towards the euarthropod condition began primarily with the limbs (so-called arthropodization) or through sclerotization of the body (arthrodization), with subsequent tagmosis

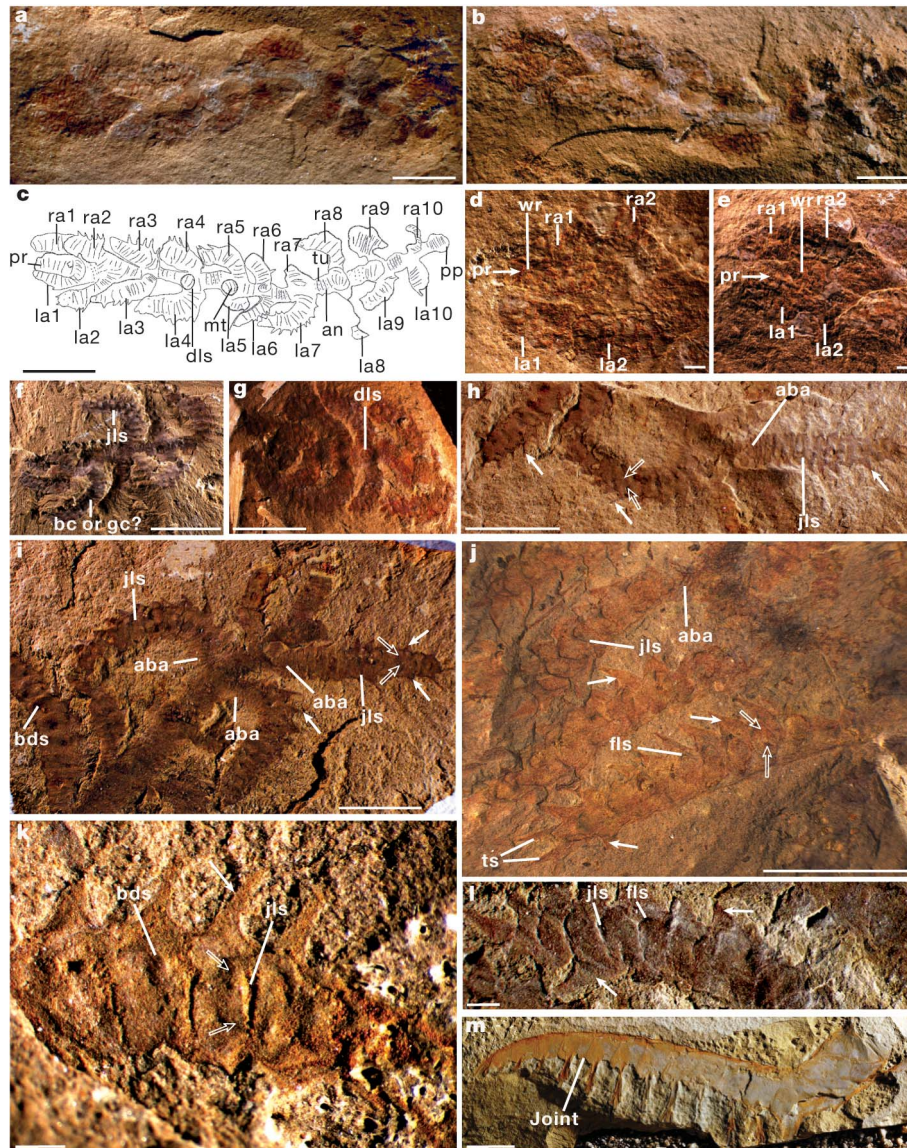


Figure 2 | Further examples of *Diania cactiformis* and comparison of the limbs to the frontal appendages of *Anomalocaris*. **a, b**, The complete specimen ELI-WD0026A,B; note the proboscis, the small protrusion and the ten pairs of appendages. **c**, Camera lucida drawing of **b**; this specimen is not twisted, thus the proboscis and the small protrusion are clearly visible. **d, e**, Enlargement of anterior part of **a** and **b**, showing the proboscis with wrinkles and the anterior two pairs of appendages, which in most cases overlapped the proboscis. **f**, ELI-WD001A, showing extension of the body cavity or gut caeca into the appendages and body cavity or gut caeca divided into several small parts by joint-like structures. **g**, S1, showing the disc-like structures on the trunk and the angle between trunk and appendages. **h–i**, ELI-EJ08-0690A,B; note the annulations at the appendage bases and the succeeding joint-like structures. **j**, S2B photographed under ethanol, appendages revealing

and cephalisation at a later date¹⁹. Some authors have specifically argued that arthrodization should have preceded arthropodization²⁰.

Our new fossils cannot resolve this question in its entirety, but they do demonstrate that appendage morphology was more diverse among Cambrian lobopodians than is sometimes realized. Even among these early faunas there may have been evolutionary grades ranging from the short, stubby but still essentially soft limbs ending in paired claws—as seen in Recent Onychophora and, for example, the Burgess Shale *Aysheaia*—through to somewhat longer, apparently still soft, but annulate ‘lobopodian’ limbs in taxa such as *Hallucigenia* and *Microdictyon*. Our *Diania* fossils further indicate the presence of lobopodians with rather robust and heavily ornamented limbs, probably finely annulated

annulations at the base followed by joint-like elements (limb articles?), the dorsal spines with round bases, the lateral spines with fold-like structures and the appendages terminating in two spines. **k**, Specimen ELI-WD008B, showing well-preserved structures resembling articulation points comparable with the condition in **k**; implying that the appendages of *Diania cactiformis* have indeed also undergone arthropodization. **bc**, body cavity; **dls**, disc-like structure; **gc**, gut caeca; **ts**, terminal spine. Unfilled white arrows point to dorsal spines, filled white arrows indicate the lateral spines. For other abbreviations see Fig. 1. Scale bars of **a–c**, **f–j** and **m** are 10 mm; those of **d, e, k** and **l** are 1 mm.

at their bases (Figs 1a, b and 2h–j), but with good evidence (see earlier) that much of the limb was divided into discrete, ring-like elements. This is evidently closer to the arthropod condition and not dissimilar to interpretations of the limb rod in ‘primitive’ arthropods and/or members of the arthropod stem-lineage such as *Fuxianhuia*²¹, or the frontal appendages of *Anomalocaris*²² (Fig. 2m) in particular.

Whether the limb ‘articles’ observed here in *Diania* were fully sclerotized and/or articulated against each other via explicit condyles like modern arthropod limbs is hard to determine from the material available although, significantly, slight projections resembling articulation points are evident (Fig. 2k, l) and reinforces their interpretation as truly jointed appendages. Nevertheless, we reiterate here that, to our

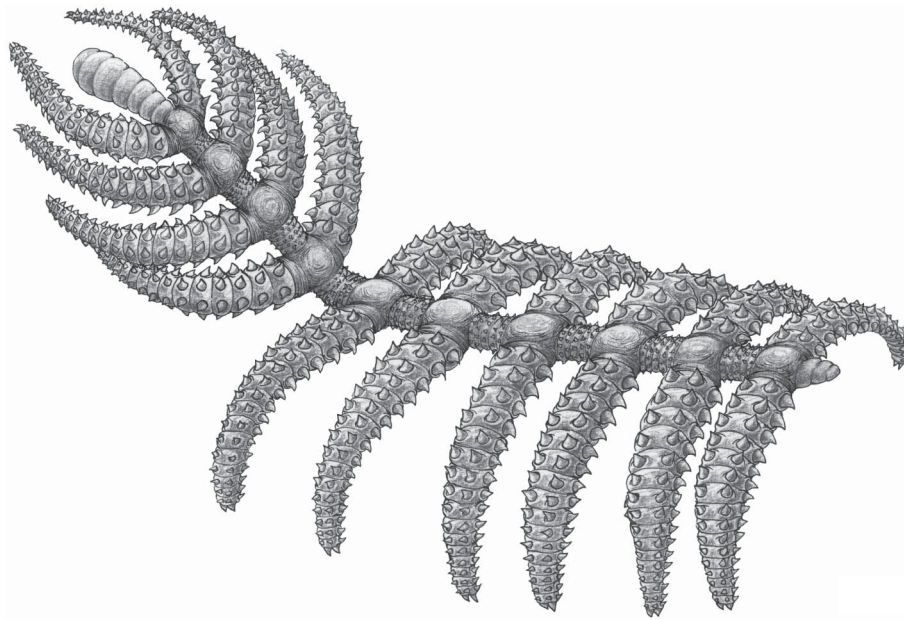


Figure 3 | Reconstruction of *Diania cactiformis* in dorsolateral view.

knowledge, *Diania* has the most robust and arthropod-like limbs found in any lobopodian until now. Given that the trunk shows little evidence for tergites or sternites (or indeed clear segments of any sort), a possible inference would be that arthropodization actually preceded arthrodization; at least within this branch of the lobopodians.

Does this 'leg-first' scenario also apply to Arthropoda in general? In part, this comes back to the problem of whether lobopodians are a clade or a (paraphyletic) grade^{10,18} in which some lineages are closer to arthropods than to modern velvet worms. Cladistic analysis of lobopodians and other stem-group arthropods (Fig. 4) does indeed resolve lobopodians as paraphyletic, with one lineage leading towards Onychophora and another towards Arthropoda. *Diania* is recovered here in the strict consensus tree as the sister-taxon of Arthropoda, on the basis of the putative apomorphy of jointed trunk appendages. Our outgroups to (*Diania* + Arthropoda) are an assemblage of fossils largely equivalent to the Dinocaridida in the sense used in refs 23 and 18, which encompasses the unusual 'gilled' lobopodians *Kerygmachela* and *Pambdelurion* as well as putative stem-group arthropods such as *Opabinia* and the Radiodonta group of *Anomalocaris* and similar forms. However, we should caution that dinocaridids, *Diania* and other potential stem-arthropods typically express mosaics of arthropod-like characters, which makes resolving a single, simple tree of arthropod origins problematic. Indeed, the position recovered here for *Diania* between Radiodonta and the ostensibly similar-looking *Schinderhannes*²⁴ is surprising. *Diania*, *Schinderhannes* and the remaining Arthropoda all share the putative apomorphy of jointed trunk appendages, and yet the trunk limbs of *Diania* resemble the frontal appendages of *Anomalocaris* and other radiodontans, which themselves lack trunk limbs entirely. If this is a secondary reduction in fossils like *Anomalocaris*, then *Diania* may in fact occupy a more basal position with respect to Radiodonta, *Schinderhannes* and Arthropoda; a scenario that would be more consistent with their fairly simple body morphology (see Supplementary Information for discussion).

Further, *Kerygmachela* and *Pambdelurion* have—compared to *Diania*—more typically lobopodian and less obviously jointed walking limbs¹⁰, but share with *Opabinia* and Radiodonta the apomorphy of lateral lobes. One possible scenario^{9,10} involves these (gill?) lobes contributing to the eventual outer branch of the arthropod biramous limb, and there is no evidence for such lobes in *Diania*. Productive areas for future research may therefore be to investigate: (1) whether leg sclerotization really did precede body sclerotization; and (2) the timing of the appearance and evolution of the endopod limb branch relative to the exopod.

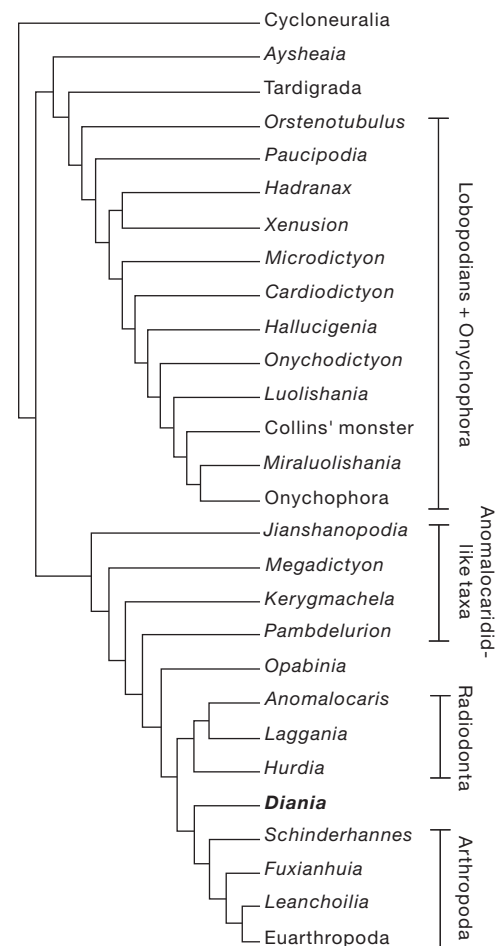


Figure 4 | Cladistic analysis of all Cambrian lobopodians and some arthropod stem group taxa. This is a strict consensus of three trees found using a branch and bound search under implied weights ($k = 2$). Tree length 130; consistency index (CI) 0.61; retention index (RI) 0.52. Tree rooted by taking Cycloneuralia as an outgroup. See Supplementary Fig. 1 and Supplementary Table 1.

Irrespective of its exact position, *Diania*, with its stout and spiny limbs attached to a slender, vermiform body, remains a highly unusual creature. It is hard to envisage it as the progenitor of any modern arthropod group, yet it may derive from a grade of lobopodian that acquired a key evolutionary innovation—and the name-giving character for Arthropoda—whereby sclerotized, jointed appendages began to fully develop.

METHODS SUMMARY

The holotype and other figured specimens with the prefix ELI are held in the Early Life Institute, Northwest University, Xi'an, China. Additional figured specimens with the prefix S are held in the collections of the Department of Earth Sciences of the Freie Universität Berlin. Most fossils were photographed with a dhs Microcam 3.3 through the ocular of a Leica Stereomicroscope MZ75; others were photographed (see figure legends for details) with a Nikon D300 specifically under 100% ethanol. Drawings were made with a camera lucida on a Leica MZ75 Stereomicroscope. The figures were prepared with Adobe Photoshop 7, Coral Draw 9 and Adobe Illustrator Artwork 13.

Received 3 October; accepted 19 November 2010.

- Hou, X.-G., Ramsköld, L. & Bergström, J. Composition and preservation of the Chengjiang fauna—a lower Cambrian soft-bodied biota. *Zool. Scr.* **20**, 395–411 (1991).
- Shu, D.-G. *et al.* Lower Cambrian vertebrates from south China. *Nature* **402**, 42–46 (1999).
- Chen, J.-Y. & Zhou, G.-Q. Biology of the Chengjiang fauna. *Bull. Natl Mus. Nat. Sci.* **10**, 11–115 (1997).
- Whittington, H. B. The lobopod animal *Aysheaia pedunculata* Walcott, Middle Cambrian, Burgess Shale, British Columbia. *Phil. Trans. R. Soc. B* **284**, 165–197 (1978).
- Hou, X.-G. & Bergström, J. Cambrian lobopodians—ancestors of extant onychophorans? *Zool. J. Linn. Soc.* **114**, 3–19 (1995).
- Ramsköld, L. & Chen, J.-Y. in *Arthropod Fossils and Phylogeny* (ed. Edgecombe, G.) 77–93 (Columbia Univ. Press, 1998).
- Delle Cave, L. & Simonetta, A. M. Notes on the morphology and taxonomic position of *Aysheaia* (Onychophora?) and of *Skania* (undetermined phylum). *Monit. Zool. Ital.* **9**, 67–81 (1975).
- Budd, G. E. Tardigrades as “stem-group arthropods”: the evidence from the Cambrian fauna. *Zool. Anz.* **240**, 265–279 (2001).
- Budd, G. E. A Cambrian gilled lobopod from Greenland. *Nature* **364**, 709–711 (1993).
- Budd, G. E. The morphology and phylogenetic significance of *Kerygmachela kierkegaardii* Budd (Buen Formation, Lower Cambrian, N Greenland). *Trans. R. Soc. Edinb. Earth Sci.* **89**, 249–290 (1999).
- Liu, J.-N., Shu, D.-G., Han, J. & Zhang, Z.-F. A rare lobopod with well-preserved eyes from Chengjiang Lagerstätte and its implications for origin of arthropods. *Chin. Sci. Bull.* **49**, 1063–1071 (2004).
- Liu, J.-N., Shu, D.-G., Han, J., Zhang, Z.-F. & Zhang, X.-L. Morpho-anatomy of the lobopod *Megadictyon cf. haikouensis* from the Early Cambrian Chengjiang Lagerstätte, South China. *Acta Zool.* **88**, 279–288 (2007).
- Liu, J.-N., Shu, D.-G., Han, J., Zhang, Z.-F. & Zhang, X.-L. A large xenusiid lobopod with complex appendages from the Chengjiang Lagerstätte (Lower Cambrian, China). *Acta Palaeontol. Pol.* **51**, 215–222 (2006).
- Dzik, J. & Krumbiegel, G. The oldest ‘onychophoran’ *Xenusion*: a link connecting phyla? *Lethaia* **22**, 169–181 (1989).
- Edgecombe, G. D. Arthropod phylogeny: an overview from the perspectives of morphology, molecular data and the fossil record. *Arthropod Struct. Dev.* **39**, 74–87 (2010).
- Maas, A., Mayer, G., Kristensen, R. M. & Waloszek, D. A Cambrian micro-lobopodian and the evolution of arthropod locomotion and reproduction. *Chin. Sci. Bull.* **52**, 3385–3392 (2007).
- Waloszek, D., Chen, J.-Y., Maas, A. & Wang, Y.-Q. Early Cambrian arthropods—new insights into arthropod head and structural evolution. *Arthropod Struct. Dev.* **34**, 189–205 (2005).
- Ma, X.-Y., Hou, X.-G. & Bergström, J. Morphology of *Luolishania longicruris* (Lower Cambrian, Chengjiang Lagerstätte, SW China) and the phylogenetic relationships within lobopodians. *Arthropod Struct. Dev.* **38**, 271–291 (2009).
- Willmer, P. *Invertebrate Relationships: Patterns in Animal Evolution* 273–296 (Cambridge Univ. Press, 1991).
- Wang, X.-Q. & Chen, J.-Y. Possible developmental mechanisms underlying the origin of crown lineages of arthropods. *Chin. Sci. Bull.* **49**, 49–53 (2004).
- Bergström, J., Hou, X.-G., Zhang, X.-G. & Clausen, S. A new view of the Cambrian arthropod *Fuxianhuia*. *GFF* **130**, 189–201 (2008).
- Hou, X.-G., Bergström, J. & Ahlberg, P. *Anomalocaris* and other large animals in the Lower Cambrian Chengjiang fauna of southwest China. *GFF* **117**, 163–183 (1995).
- Collins, D. The “evolution” of *Anomalocaris* and its classification in the arthropod class Dinocarida (nov.) and order Radiodonta (nov.). *J. Palaeont.* **70**, 280–293 (1996).
- Kühl, G., Briggs, D. E. G. & Rust, J. A Great-Appendage arthropod with a radial mouth from the Lower Devonian Hunsrück Slate, Germany. *Science* **323**, 771–773 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by a Research Scholarship from the Alexander von Humboldt Foundation at the Freie Universität Berlin, the National Science Foundation of China (grants 40802011 and 40830208) and the MOST Special Fund from the State Key Laboratory of Continental Dynamics, Northwest University (to J.L.), and a DFG grant within the Forschergruppe 736 (to H.K. and M.S.). We thank M. G. Chi for the reconstruction, J. Evers and M. Y. Sun for photography and table, and local workers for their strenuous field work.

Author Contributions J.L. collected most of the fossils, described them and wrote the paper with the other authors; M.S. collected material and was involved in the phylogenetic analysis; J.A.D. contributed to the discussion; H.K. and D.S. were involved in technical aspects of the analysis; Q.O. provided three specimens. J.H. and Z.Z. contributed to fieldwork; X.Z. was involved in the analysis.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.L. (liujianni@126.com).

Consequences of climate change on the tree of life in Europe

Wilfried Thuiller¹, Sébastien Lavergne¹, Cristina Roquet¹, Isabelle Boulangeat¹, Bruno Lafourcade¹ & Miguel B. Araujo^{2,3}

Many species are projected to become vulnerable to twenty-first-century climate changes^{1,2}, with consequent effects on the tree of life. If losses were not randomly distributed across the tree of life, climate change could lead to a disproportionate loss of evolutionary history^{3–5}. Here we estimate the consequences of climate change on the phylogenetic diversities of plant, bird and mammal assemblages across Europe. Using a consensus across ensembles of forecasts for 2020, 2050 and 2080 and high-resolution phylogenetic trees, we show that species vulnerability to climate change clusters weakly across phylogenies. Such phylogenetic signal in species vulnerabilities does not lead to higher loss of evolutionary history than expected with a model of random extinctions. This is because vulnerable species have neither fewer nor closer relatives than the remaining clades. Reductions in phylogenetic diversity will be greater in southern Europe, and gains are expected in regions of high latitude or altitude. However, losses will not be offset by gains and the tree of life faces a trend towards homogenization across the continent.

Evidence is accumulating that ongoing climate changes already affect living organisms⁶. Although species have been continuously exposed to climate changes throughout their evolutionary histories⁷, there is concern that the pace of current changes poses unprecedented challenges for many species⁸. Evidence exists that certain clades are more vulnerable to anthropogenic pressures than others⁹, causing portions of Earth's evolutionary history to be disproportionately altered. Here we investigate whether climate change could have a similar effect.

Climatic tolerances vary across species, causing some species to be more vulnerable to climate change than others¹⁰. Because climate tolerances are not randomly distributed across phylogenies¹¹, species sensitivities to climate change are expected to be clustered along the phylogeny. It follows that if vulnerable species are closely related, shared internal branches of the tree of life have higher risks of collapsing¹² (Supplementary Fig. 1). In contrast, overdispersed extinctions in the phylogeny would mitigate the loss of internal edges⁹. However, phylogenetically clumped extinctions should cause disproportionate loss of phylogenetic diversity only when extinctions are in clades with long edges overall⁵ (Supplementary Fig. 1, scenario A). In contrast,

extinctions clumped in recent—often widespread and fast-evolving—clades, with short branches (Supplementary Fig. 1, scenario B), should have impacts no different, or even smaller, than under a model of random extinctions⁹ (Supplementary Fig. 1, scenario C).

We report a comprehensive analysis of the impacts of climate change on the European tree of life (see Supplementary Fig. 2 for a workflow of the analysis). First, we test whether changes in suitable climate (consensus projections extracted from several species distribution models and high-resolution climate model projections) among 1,280 plant, 340 bird and 140 mammal European species show a significant phylogenetic signal. Second, using contractions of suitable climate as proxies for extinction risk, we test whether losses would cause greater decline in phylogenetic diversity than expected under a null model of extinction. Third, we ask whether expected changes in phylogenetic diversity are spatially structured. To this end, we map the current and future phylogenetic diversity of each study group and investigate changes in regional phylogenetic diversity and its spatial turnover¹³ across Europe.

Changes in suitable climate among European species were generally more similar between closely related species, but the strength of the phylogenetic signal was weak (Fig. 1, Supplementary Table 1 and Supplementary Figs 3 and 4). This pattern of clustering appeared consistently across the high-resolution phylogenies, emission scenarios and periods studied (Supplementary Table 1 and Supplementary Figs 3 and 4). For example, among plants, most suitable climate for *Draba* (whitlow grasses) and *Arabis* (rockcress) species contracted (72% of *Draba* species contracted by >30% and 50% of *Arabis* species contracted by >50%). Among birds, most *Tringa* (shanks and tattlers) and *Numenius* (curlew) species had suitable climate contractions of >30%, whereas most *Ardeidae* (herons and egrets) expanded their ranges. Mammals were generally less vulnerable to climate change, although more than half of the *Sorex* (long-tailed shrews) species could lose more than 30% of their suitable climate (Fig. 1 and Supplementary Fig. 5).

Reductions in phylogenetic diversity arising from climate change were not greater than expected under a model of random extinctions

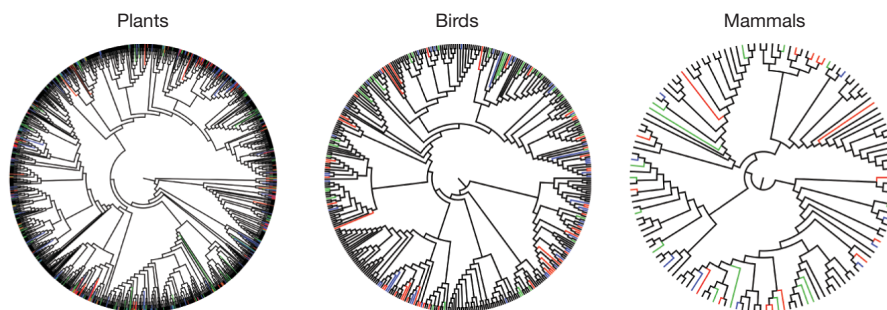


Figure 1 | Changes in suitable climate (A1FI scenario for 2080) mapped onto the phylogeny of European plants, birds and mammals. Black edges indicate positive changes in suitable climate (CSC), that is, range expansion.

Other colours indicate negative changes in suitable climate (range contraction): green for CSC > -15%, blue for CSC > -30% and red for CSC > -50%.

¹Laboratoire d'Ecologie Alpine, UMR CNRS 5553, Université Joseph Fourier, BP 53, FR-38041 Grenoble Cedex 9, France. ²Department of Biodiversity and Evolutionary Biology, National Museum of Natural Sciences, CSIC, Calle Gutierrez Abascal, 2, 28006, Madrid, Spain. ³Rui Nabeiro Biodiversity Chair, CIBIO, University of Évora, Largo dos Colegiais, 7000 Évora, Portugal.

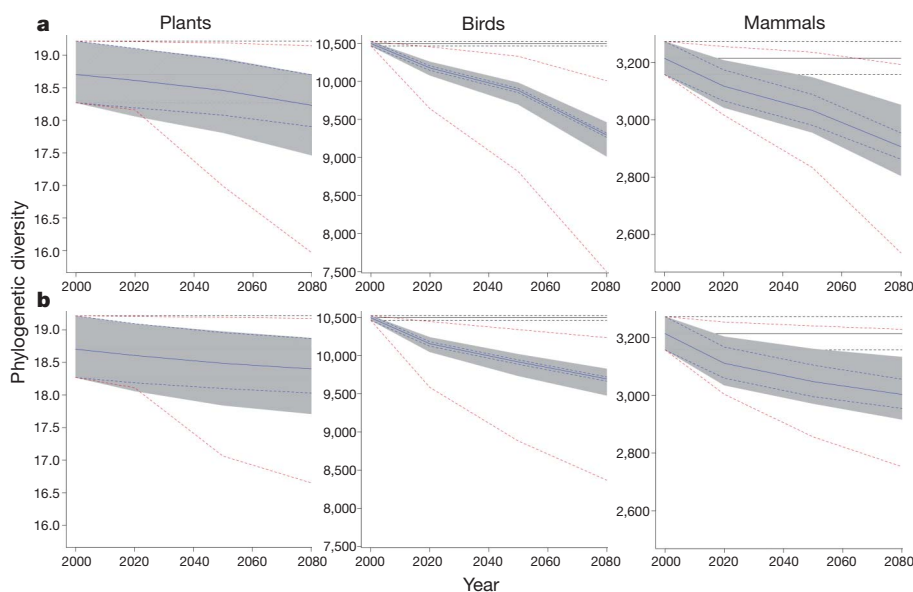


Figure 2 | Changes in phylogenetic diversity versus scenarios of random extinction for plants, birds and mammals. **a**, Emission scenario A1FI; **b**, emission scenario B1. Black solid and dashed lines depict the median, maximum and minimum current phylogenetic diversities across the sample of trees. Blue solid and dashed lines represent the median, maximum and minimum projected phylogenetic diversities due to range change across the

sample of trees. The grey area is the quantile range of projected phylogenetic diversity due to range contraction (from 2020 onward), randomly scattered across the sample of trees. The red lines are the remaining phylogenetic diversity when the risk of extinction is positively (lower line) or negatively (upper line) related to the evolutionary distinctiveness of the taxa.

across the phylogeny (Fig. 2 and Supplementary Fig. 6). That is, species whose suitable climates contract are not clustered in species-poor clades with long branches (where projected future phylogenetic diversity should be lower than under random species losses), nor are they clustered on the shortest branches (where projected future phylogenetic diversity should be higher than under random species losses). Instead they are clustered across both long and short edges of the tree. Again, these results were consistent across the three clades studied, the sampled trees, the four emission scenarios and the time slices 2050 and 2080. Sensitivity analysis across climate model projections and species distribution models revealed that results were also not affected by variability across models (Supplementary Fig. 7). For plants at risk, reductions in phylogenetic diversity were relatively low: approximately 2.7% in the worst-case scenario (A1FI) and for the time slice with greater impacts, that is, 2080 (current phylogenetic diversity, 18.7; future phylogenetic diversity, 18.2). Reductions of the same order of magnitude were recorded for phylogenetic diversity in birds (11.5% reduction) and mammals (9.6% reduction).

Spatial analyses of projected phylogenetic diversity revealed a marked homogenization of phylogenetic diversity across Europe, with strong reductions of spatial turnover (approximately –34%, –32% and –30% for plants, birds and mammals, respectively) following climate change. However, regions are not all equally affected by climate changes (Fig. 3). At present, the spatial distribution of plant phylogenetic diversity separates the northern Iberian Peninsula, Italy and France, with high total phylogenetic diversity, from northern European countries and the Alps, with low total phylogenetic diversity (Fig. 3). Similar patterns are found for birds, with higher phylogenetic diversity in the Mediterranean basin than in the rest of Europe, and mammals, with higher phylogenetic diversity in the southeast of Europe than in the northeast. Increases in phylogenetic diversity in northern Europe and in the Alps are accompanied by consistent decreases in phylogenetic diversity in southern Europe, causing a general reduction in the spatial variation of phylogenetic diversity (Fig. 3). These predicted spatial changes in phylogenetic diversity are not entirely due to changes in species richness (Supplementary Fig. 8). Indeed, regions such as Scandinavia (for plants), Germany (for birds) and Sweden (for mammals) are projected

to have more phylogenetic diversity in the future than is expected on the basis of their future projected species richness (Supplementary Fig. 8). This can be explained by northern biotas being more likely to sample species that are less phylogenetically redundant in the future. The projected reduction of phylogenetic diversity in southern Europe could, however, be alleviated by species immigration from Northern Africa, especially if immigrant species belong to phylogenetic clades that are not present in the northern Mediterranean basin.

Linking phylogenetic and biogeographic information can help identify regions of past production and present maintenance of biodiversity (so-called cradles and graves)¹⁴, but projecting them into the future is challenging. Our study addresses this challenge and presents a unique large-scale assessment of the potential impacts of climate change on the evolutionary history of plants, birds and mammals. Although our assessment integrates uncertainty in phylogenetic reconstructions, it should be noted that projected changes in evolutionary history are also inevitably sensitive to the species distribution and climate models used. To address this problem, we have used cutting-edge bioclimatic ensemble forecasting methodologies. Because high-resolution climate projections have large uncertainties (owing to the difficulty of simulating local climates and the inaccuracy of interpolation techniques¹⁵), we have included three well-known global change models encompassing a large variation in future climate and shown that our results were robust to this variation. Accordingly, we show that although different metrics of species vulnerability to climate change tend not to be randomly distributed with regards to the tree of life, the loss of European phylogenetic diversity is not greater than expected under a model of random extinctions⁴. The International Union for Conservation of Nature Red List of Threatened Species provides clear evidence that extinction risk is selective among particular groups such as the amphibians, birds and mammals¹⁶. The prevalence of threatened species differs significantly among these groups and among the families and orders within each group^{9,16}. The fossil record also provides evidence of phylogenetically clustered extinctions, although the previous five mass extinctions provide examples of less extreme selectivity^{17,18}. Our projections do not predict a large drop in total phylogenetic diversity, but they do suggest a future restructuring of the spatial distribution of

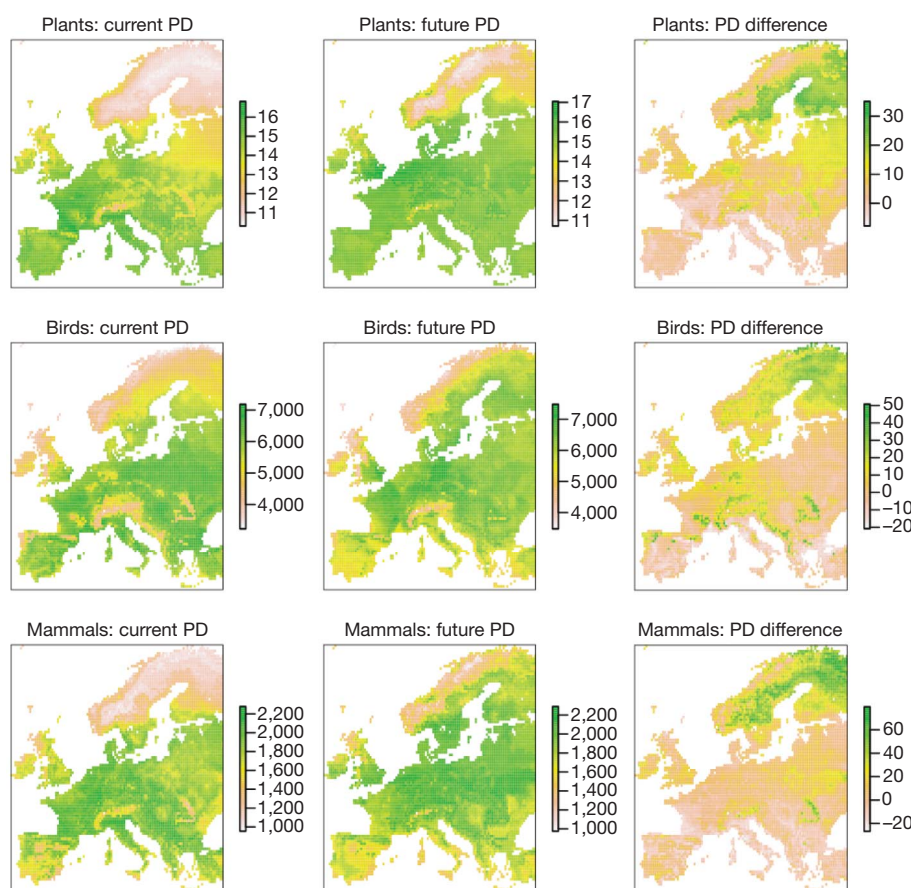


Figure 3 | Map of current and future phylogenetic diversities (A1FI scenario for 2080) and their relative differences for the three species groups. Maps represent average phylogenetic diversity (PD; colour scale) across the sample of 100 phylogenetic trees used for each study group.

the tree of life. Phylogenetic diversity over Europe will be homogenized owing to the reshuffling of species assemblages and the migration of the tree of life into higher latitudes and elevations.

METHODS SUMMARY

We modelled changes in suitable climate for 1,280 European plants, 140 mammals and 340 birds, using six species distribution models¹⁹, three 10' high-resolution global climate models (HadCM3, CSIRO2, CGCM2¹⁵ from CRU-TYN-SC-1.0) and four emission scenarios (A1FI, A2, B1 and B2²⁰), over the periods 1961–1990, 1991–2020, 2021–2050 and 2051–2080. Given the low variability across species distribution models and global climate models (Supplementary Information and Supplementary Fig. 7), we kept the median consensus projection for each scenario and time slice²¹.

Species vulnerability was estimated as the change in the total area of suitable climate assuming unlimited dispersal (CSC, ranging from -100 to >100). Another measure ('loss' of suitable climate (LSC), ranging from -100 to 0) gave similar results (Supplementary Figs 4 and 5).

We developed mega-phylogenies by mixing supertree and supermatrix approaches²² with sequences from GenBank. Optimal alignment for each region was obtained with four algorithms, and by choosing and deparating the best alignment with TRIMAL. Phylogenetic analyses were conducted using maximum-likelihood estimation by constraining heuristic searches with a family-level supertree. We retained 100 maximum-likelihood trees for plants and birds. For mammals, we used 100 phylogenetic trees based on ref. 23 with polytomies randomly resolved.

Estimation and test of phylogenetic signal in species vulnerability was performed using Abouheif's test²⁴, Blomberg's K (ref. 25) and Pagel's λ (ref. 26).

To estimate the potential impacts of climate change on the tree of life, we used CSC as a surrogate for probability of extinction, $p(\text{ext})$, and calculated future phylogenetic diversity by weighting the probability of an edge persisting as $1 - p(\text{ext})$ of all its descendent species²⁷. The null model expectation was extracted by randomizing $p(\text{ext})$ across the tips and recalculating phylogenetic diversity.

Spatial distribution of phylogenetic diversity was estimated on the pixel basis with species projected to be present. Spatial turnover was measured as the total projected phylogenetic diversity minus the mean local phylogenetic diversity over Europe¹³.

Received 7 June; accepted 29 November 2010.

Published online 16 February 2011.

1. Thomas, C. D. *et al.* Extinction risk from climate change. *Nature* **427**, 145–148 (2004).
2. Thuiller, W., Lavorel, S., Araújo, M. B., Sykes, M. T. & Prentice, I. C. Climate change threats to plant diversity in Europe. *Proc. Natl Acad. Sci. USA* **102**, 8245–8250 (2005).
3. Mace, G. M., Gittleman, J. L. & Purvis, A. Preserving the tree of life. *Science* **300**, 1707–1709 (2003).
4. Nee, S. & May, R. M. Extinction and the loss of evolutionary history. *Science* **288**, 328–330 (1997).
5. Heard, S. B. & Mooers, A. O. Phylogenetically patterned speciation rates and extinction risks change the loss of evolutionary history during extinctions. *Proc. R. Soc. Lond. B* **267**, 613–620 (2000).
6. Parmesan, C. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Syst.* **37**, 637–669 (2006).
7. Meyers, L. A. & Bull, J. J. Fighting change with change: adaptive variation in an uncertain world. *Trends Ecol. Evol.* **17**, 551–557 (2002).
8. Thuiller, W. Climate change and the ecologist. *Nature* **448**, 550–552 (2007).
9. Purvis, A. Phylogenetic approaches to the study of extinction. *Annu. Rev. Ecol. Syst.* **39**, 301–319 (2008).
10. Thuiller, W., Lavorel, S. & Araújo, M. B. Niche properties and geographical extent as predictors of species sensitivity to climate change. *Glob. Ecol. Biogeogr.* **14**, 347–357 (2005).
11. Prinzing, A., Durka, W., Klotz, S. & Brandl, R. The niche of higher plants: evidence for phylogenetic conservatism. *Proc. R. Soc. Lond. B* **268**, 2383–2389 (2001).
12. McKinney, M. L. Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annu. Rev. Ecol. Syst.* **28**, 495–516 (1997).
13. Lande, R. Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* **76**, 5–13 (1996).
14. Chown, S. L. & Gaston, K. J. Areas, cradles and museums: the latitudinal gradient in species richness. *Trends Ecol. Evol.* **15**, 311–315 (2000).

15. Mitchell, T. D. & Jones, P. D. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.* **25**, 693–712 (2005).
16. Baillie, J. E. M., Hilton-Taylor, C. & Stuart, S. N. (eds) *2004 IUCN Red List of Threatened Species. A Global Species Assessment* (International Union for Conservation of Nature, 2004).
17. Bambach, R. K. Phanerozoic biodiversity mass extinctions. *Annu. Rev. Earth Planet. Sci.* **34**, 127–155 (2006).
18. Koch, P. L. & Barnosky, A. D. Late Quaternary extinctions: state of the debate. *Annu. Rev. Ecol. Syst.* **37**, 215–250 (2006).
19. Thuiller, W., Lafourcade, B., Engler, R. & Araujo, M. B. BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography* **32**, 369–373 (2009).
20. Nakicenovic, N. & Swart, R. *Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change* 570 (Cambridge Univ. Press, 2000).
21. Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R. K. & Thuiller, W. Evaluation of consensus methods in predictive species distribution modelling. *Divers. Distrib.* **15**, 59–69 (2009).
22. Smith, S. A., Beaulieu, J. M. & Donoghue, M. J. Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. *BMC Evol. Biol.* **9**, 37–48 (2009).
23. Fritz, S. A., Bininda-Emonds, O. R. P. & Purvis, A. Geographical variation in predictors of mammalian extinction risk: big is bad, but only in the tropics. *Ecol. Lett.* **12**, 538–549 (2009).
24. Pavoine, S., Ollier, S., Pontier, D. & Chessel, D. Testing for phylogenetic signal in phenotypic traits: new matrices of phylogenetic proximities. *Theor. Popul. Biol.* **73**, 79–91 (2008).
25. Blomberg, S. P., Garland, T. & Ives, A. R. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution* **57**, 717–745 (2003).
26. Pagel, M. Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884 (1999).
27. Witting, L. & Loeschke, V. The optimization of biodiversity conservation. *Biol. Conserv.* **71**, 205–207 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This research was funded by the EU ECOCHANGE (GOCE-CT-2007-036866) and DIVERSITALP (ANR-2007-BDIV-014) projects. C.R. was supported by a grant from Fundación Ramón Areces. We thank P. Pearman and A. Mooers for comments on earlier drafts. Computations were performed using the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>), supported by the Rhône-Alpes region (GRANT CPER07-13 CIRA).

Author Contributions W.T. and S.L. designed the study, C.R. built the plant and bird phylogenies, I.B. and B.L. helped with R code writing, and W.T. performed all the analyses. W.T., S.L. and M.B.A. wrote the manuscript with substantial contributions from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to W.T. (wilfried.thuiller@ujf-grenoble.fr).

Synaptic potentiation onto habenula neurons in the learned helplessness model of depression

Bo Li^{1,2*}, Joaquin Piriz^{1*}, Martine Mirrione^{2,3*}, ChiHye Chung^{1*}, Christophe D. Proulx¹, Daniela Schulz³, Fritz Henn^{2,3} & Roberto Malinow¹

The cellular basis of depressive disorders is poorly understood¹. Recent studies in monkeys indicate that neurons in the lateral habenula (LHb), a nucleus that mediates communication between forebrain and midbrain structures, can increase their activity when an animal fails to receive an expected positive reward or receives a stimulus that predicts aversive conditions (that is, disappointment or anticipation of a negative outcome)^{2–4}. LHb neurons project to, and modulate, dopamine-rich regions, such as the ventral tegmental area (VTA)^{2,5}, that control reward-seeking behaviour⁶ and participate in depressive disorders⁷. Here we show that in two learned helplessness models of depression, excitatory synapses onto LHb neurons projecting to the VTA are potentiated. Synaptic potentiation correlates with an animal's helplessness behaviour and is due to an enhanced presynaptic release probability. Depleting transmitter release by repeated electrical stimulation of LHb afferents, using a protocol that can be effective for patients who are depressed^{8,9}, markedly suppresses synaptic drive onto VTA-projecting LHb neurons in brain slices and can significantly reduce learned helplessness behaviour in rats. Our results indicate that increased presynaptic action onto LHb neurons contributes to the rodent learned helplessness model of depression.

To study the cellular basis of behavioural depression, we examined the synaptic circuitry in the LHb of rats showing learned helplessness (Fig. 1), a model of depression whereby animals show reduced escape from escapable foot shock¹⁰. We used two well-established animal models: acute learned helplessness (aLH), which is induced by subjecting rats to periods of inescapable and unpredictable shock¹⁰; and congenital learned helplessness (cLH), which is present in a strain of rats produced by selective breeding of animals that show the greatest amount of aLH^{11,12}. In addition to showing reduced escape from escapable foot shock, animals with cLH and aLH also showed greater immobility in the forced swim test, another widely used animal model for depression¹³, than control animals (Fig. 2a, b).

We examined transmission onto LHb neurons. These neurons receive major inputs from numerous brain regions involved in stress response (such as the entopeduncular nucleus, lateral hypothalamus, lateral preoptic area, medial prefrontal cortex and the bed nucleus of the stria terminalis)¹⁴ (Supplementary Fig. 1) and can control dopaminergic function in the midbrain⁴. We wished to determine whether synaptic transmission onto LHb neurons is different in animals with learned helplessness and normal animals. To record selectively from LHb neurons that can regulate the activity of dopamine-producing neurons, we injected a retrograde tracer, cholera toxin conjugated to the dye Alexa Fluor 488, into the VTA *in vivo*. Two to three days later, we prepared brain slices that contained the LHb. A minority of neurons in the LHb were fluorescent, which indicated their projection to the VTA (Supplementary Fig. 2a, b). Notably, the LHb neurons that project to the VTA and to the rostromedial tegmental nucleus—a newly identified GABA (γ -aminobutyric acid)-producing, inhibitory relay station

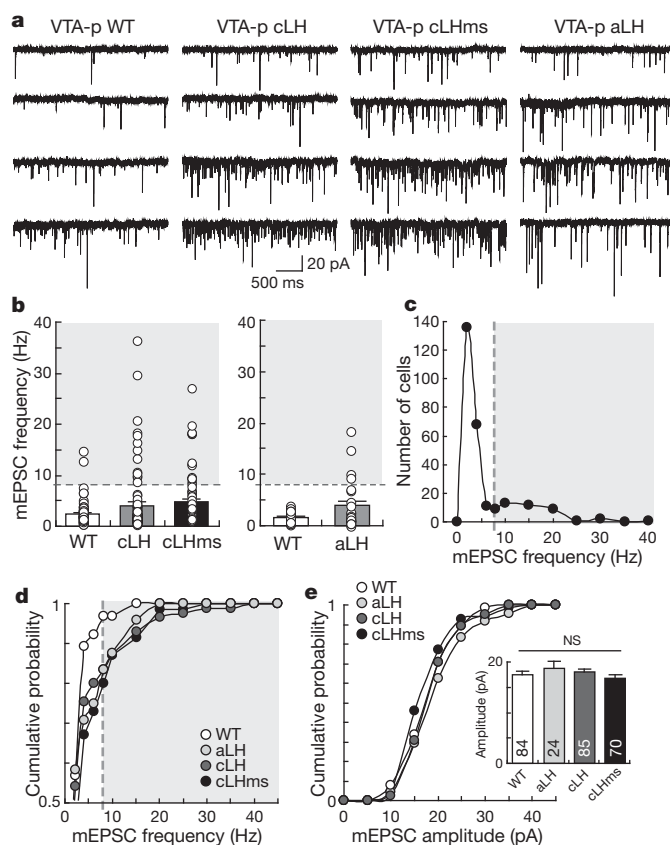


Figure 1 | Increased excitatory synaptic transmission onto VTA-projecting LHb neurons in the learned helplessness models of depression. **a**, Examples of mEPSCs recorded from VTA-projecting LHb neurons (VTA-p) from wild-type control (WT) animals and animals with cLH, cLHms or aLH. **b**, Means (histogram bars) and individual recordings (open circles) of mEPSC frequency from VTA-projecting LHb neurons in various groups of animals. Left: WT, 2.4 ± 0.3 , $n = 65$ (6 animals); cLH, 4.0 ± 0.7 , $n = 85$ (8 animals), $P < 0.05$ bootstrap method; cLHms, 4.7 ± 0.7 , $n = 70$ (8 animals), $P < 0.001$ bootstrap. Right: WT, 1.8 ± 0.2 , $n = 19$ (4 animals); aLH, 3.7 ± 0.8 , $n = 23$ (4 animals), $P < 0.05$ bootstrap. Results are presented as mean \pm s.e.m. P values are compared with WT group. Shaded region beyond dashed lines indicates high-frequency mEPSCs. **c**, Frequency distribution of mEPSC frequencies of all cells recorded ($n = 263$ from 34 animals) showed bimodal distribution. **d**, The cumulative probability of mEPSC frequency of VTA-projecting LHb neurons in different groups of animals ($P < 0.05$, Kolmogorov–Smirnov (K–S) test comparing WT with any other group). **e**, The amplitude of mEPSCs did not differ among the four animal groups ($P > 0.3$, bootstrap). Inset shows mean \pm s.e.m. n , number of cells; NS, not significant.

¹Center for Neural Circuits and Behavior, Departments of Neuroscience and Biological Sciences, 9500 Gilman Drive # 0634, University of California at San Diego, La Jolla, California 92093, USA. ²Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. ³Brookhaven National Laboratory, Upton, New York 11973, USA.

*These authors contributed equally to this work.

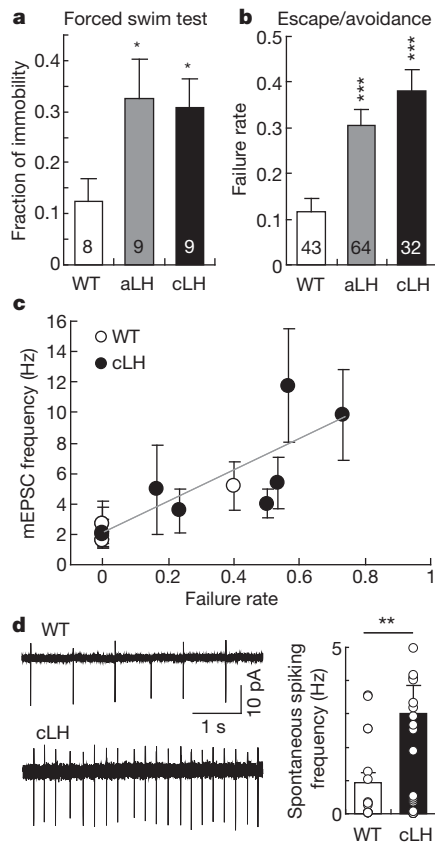


Figure 2 | Enhanced synaptic transmission onto VTA-projecting LHB neurons correlates with helpless behaviour of individual animals.

a, b, Animals with aLH or cLH show behavioural deficits in the forced swim test and in the escape/avoidance test. **a**, Fraction of immobile time over the 5 min of the forced swim test: WT, 0.12 ± 0.05 ; aLH, 0.33 ± 0.08 ; and cLH, 0.31 ± 0.05 . Results are mean \pm s.e.m.; n (number of animals) indicated in histogram bars; *, $P < 0.05$, Kruskal–Wallis test. **b**, Rate of failure to escape during 30 trials of escapable foot shock: WT, 0.12 ± 0.03 ; aLH, 0.31 ± 0.03 ; and cLH, 0.38 ± 0.05 . Results are mean \pm s.e.m.; n (number of animals) is indicated in the histogram bars; ****P* < 0.001, $F_{(2,136)} = 11.57$, one-way ANOVA. **c**, The mean frequency of mEPSCs onto VTA-projecting LHB neurons correlates with an animal's helpless behaviour, measured as the fraction of sessions in which animals failed to escape ($R^2 = 0.69$, grey line; $P < 0.001$ by a linear regression; $n = 13$ animals; $n \geq 5$ cells for each animal). Error bars, s.e.m. **d**, The spontaneous spiking rate measured in a cell-attached configuration was higher in animals with cLH than in WT control animals: left, an example; right, histogram shows mean \pm s.e.m., and open circles are data from individual cells (WT, 0.92 ± 0.32 , $n = 17$; cLH, 3.03 ± 0.82 ; $n = 25$, **, $P < 0.01$, bootstrap).

between the LHB and the VTA^{15,16}—are largely non-overlapping populations (Supplementary Fig. 2c), indicating that we would be able to selectively target LHB neurons that directly project to the VTA. LHB neurons projecting to the VTA were glutamatergic, as indicated by their co-localization with the glutamate transporter EAAC1 (also known as SLC1A1) and lack of GABAergic marker expression (Supplementary Fig. 3).

We performed whole-cell patch-clamp recordings on VTA-projecting neurons in acute parasagittal brain slices from rats that were wild-type control, had aLH or cLH (naive), or had cLH and had been exposed to mild stress (cLHms; see ‘Behavioural paradigms’ in the Methods section). We examined miniature excitatory postsynaptic currents (mEPSCs) (in the presence of tetrodotoxin to block action potentials and picrotoxin to block GABA_A-mediated synaptic currents), which were mediated by AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid)-type glutamate receptors (Supplementary Fig. 4a) and represent responses from individual synapses onto the cells studied.

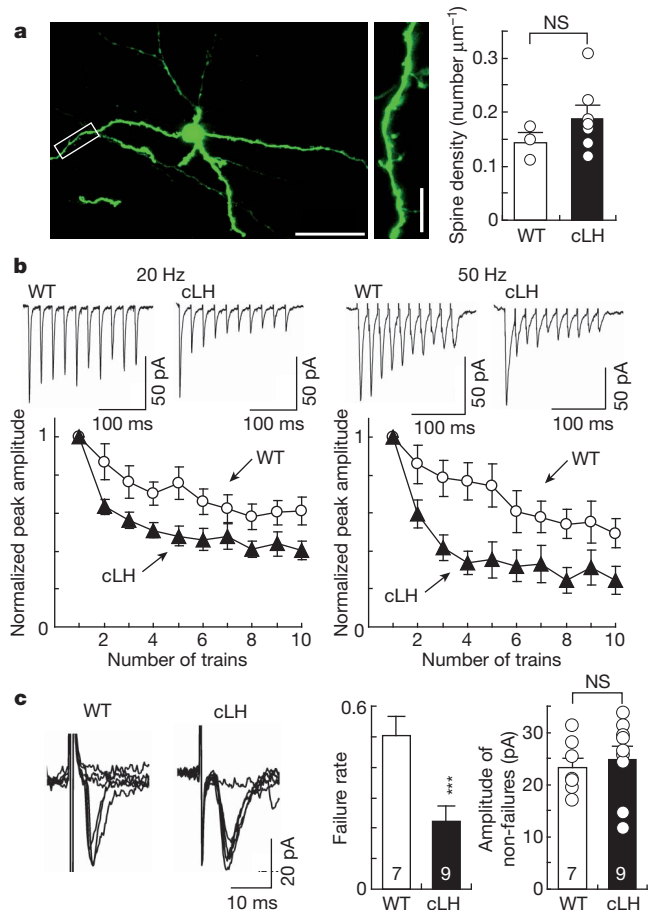


Figure 3 | Presynaptic mechanism underlying the increase in excitatory synaptic transmission onto VTA-projecting LHB neurons in helpless animals.

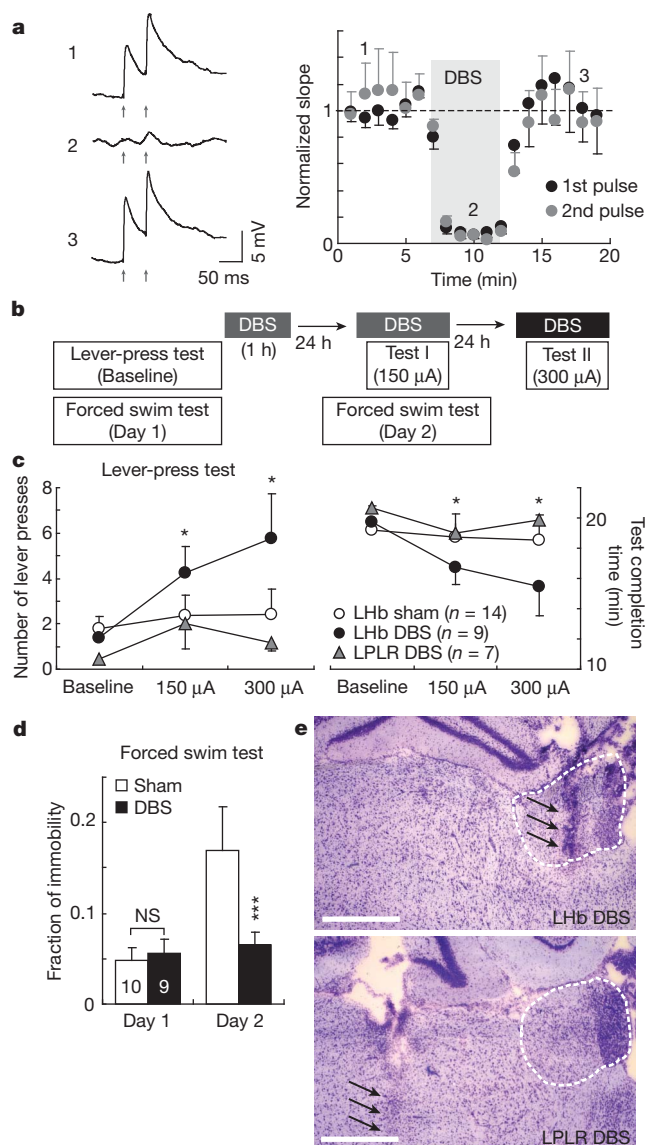
a, Two-photon laser scanning images of a VTA-projecting LHB neuron. The neuron was labelled by *in vivo* injection into the VTA of a herpes simplex virus that expresses green fluorescent protein and is transported in a retrograde manner. The neuron is shown at low magnification (left) and high magnification (centre, image of area demarcated by white rectangle in left panel). Scale bars, 50 μ m (left) and 10 μ m (centre). Right, dendritic spine density on VTA-projecting LHB neurons of WT animals or animals with cLH: WT, 0.15 ± 0.02 , 3 cells, 1,125.3 μ m total dendritic length; and cLH, 0.18 ± 0.02 , 7 cells, 1,002.4 μ m total dendritic length; $P > 0.1$, Student's *t*-test. **b**, Top, evoked EPSCs onto VTA-projecting LHB neurons of WT animals or animals with cLH in response to stimulus trains (20 Hz or 50 Hz). Bottom, plot of peak EPSCs normalized to first EPSC (WT, 20 Hz $n = 10$, 50 Hz $n = 11$; and cLH, 20 Hz $n = 14$, 50 Hz $n = 13$). Compared with WT animals, animals with cLH showed a faster synaptic depression (20 Hz, $F_{(9,198)} = 2.32$, $P = 0.02$; 50 Hz, $F_{(9,198)} = 3.83$, $P < 0.001$; one-way ANOVA with repeated measures) and more extensive synaptic depression (20 Hz, $F_{(1,22)} = 6.62$, $P = 0.02$; 50 Hz, $F_{(1,22)} = 7.15$, $P = 0.01$; one-way ANOVA with repeated measures). **c**, Minimally evoked EPSCs onto VTA-projecting LHB neurons show more failures in WT control animals than in animals with cLH (WT, 0.5 ± 0.1 , $n = 7$; cLH, 0.2 ± 0.1 , $n = 9$; ****P* < 0.001, Student's *t*-test). Mean amplitude of successful trials (right) (WT, 23.2 ± 2 , $n = 7$; cLH: 24.8 ± 2.5 , $n = 9$; $P > 0.6$, Student's *t*-test). **a–c**, Error bars, s.e.m.

The mean frequency of mEPSCs recorded from VTA-projecting LHB neurons of rats with learned helplessness—aLH (3.7 ± 0.8 Hz (mean \pm s.e.m.), $n = 23$), cLH (4.0 ± 0.7 Hz, $n = 85$) and cLHms (4.7 ± 0.7 Hz, $n = 70$)—was higher than that of wild-type controls (2.3 ± 0.2 Hz, $n = 84$; $F_{(3,251)} = 3.1$, $P < 0.03$ comparing the wild-type control group with any other group, analysis of variance (ANOVA)) (Fig. 1a, b). In general, the distribution of mEPSC frequencies recorded across different cells in all groups had a bimodal distribution (Fig. 1c). Notably, the prevalence of neurons with high-frequency mEPSCs (> 8 Hz; Fig. 1b–d, shaded region) was significantly higher

in rats with aLH (17%), cLH (14%) and cLHms (20%) than in wild-type control rats (2%; $P < 0.01$ comparing the wild-type control group with any other group, χ^2 test).

To determine whether the observed excitatory synaptic potentiation was quantitatively correlated to an animal's helpless behaviour, we first tested animals (either wild type or with cLH) by using an escape avoidance task and subsequently prepared brain slices and made recordings from the VTA-projecting neurons. For each animal, we recorded from at least five cells and plotted the mean mEPSC frequency against the animal's helpless behaviour (as measured by the fraction of trials in which the animal failed to escape from an escapable 10-s foot shock; see also Fig. 2b). The significant correlation ($R^2 = 0.69$, $F_{(1,11)} = 24.85$, $P < 0.001$, $n = 13$ for all animals; and $R^2 = 0.64$, $F_{(1,6)} = 10.7$, $P < 0.05$, $n = 8$ for animals with cLH) (Fig. 2c) indicates that the potentiation of excitatory transmission onto VTA-projecting LHB neurons is linked with an individual animal's helpless behaviour.

To examine the output of VTA-projecting neurons, we measured their spontaneous action potentials, which were more frequent in animals with cLH than in wild-type controls (Fig. 2d). We observed no differences among the various groups with respect to the amplitude of mEPSCs (Fig. 1e) or the frequency or amplitude of miniature inhibitory postsynaptic currents (Supplementary Fig. 4b, c). These results indicate that the excitatory synaptic input onto LHB neurons that project to the VTA is potentiated in the learned helplessness model.



The enhanced mEPSC frequency could result from an increase in either the number of synapses or the probability of presynaptic neurotransmitter release. To distinguish between these possibilities, we first measured the density of dendritic spines, which are the sites of excitatory synapses, on the dendrites of VTA-projecting LHB neurons. There was no significant difference in dendritic spine density between wild-type controls and animals with cLH (Fig. 3a), and there was no obvious difference in the patterns of dendritic branching between the two groups (data not shown), suggesting that there was no major difference in the number of synapses between wild-type control animals and those with cLH.

To determine whether there is a change in the efficacy of presynaptic neurotransmitter release, we examined evoked transmission. Synaptic transmission onto LHB neurons (elicited by placing a stimulating electrode in the LHB) showed distinct properties: the evoked excitatory synaptic response had a very small NMDA (*N*-methyl-D-aspartate) receptor component (Supplementary Fig. 4d), and the AMPA receptor component showed strong inward rectification (Supplementary Fig. 4e). To probe presynaptic function, we evoked transmission with high-frequency stimulation trains (ten stimuli delivered at 20 Hz or 50 Hz). The decrease in the amplitude of EPSCs in response to successive pulses during a train of stimuli reflects presynaptic vesicle depletion; more depletion correlates with a higher release probability¹⁷. VTA-projecting LHB neurons of animals with cLH showed a faster synaptic depression (at 20 Hz $F_{(9,198)} = 2.32$, $P < 0.05$; at 50 Hz $F_{(9,198)} = 3.83$, $P < 0.001$, one-way ANOVA with repeated measures) and a more extensive synaptic depression (at 20 Hz $F_{(1,22)} = 6.62$, $P < 0.05$; at 50 Hz $F_{(1,22)} = 7.15$, $P < 0.05$, one-way ANOVA with repeated measures) than those in wild-type control animals (Fig. 3b). Furthermore, with minimal stimulation, which is designed to activate few synapses (as indicated by the amplitude of non-failure responses, which is similar to the mEPSC amplitude; Fig. 3c), we measured synaptic transmission failure rate. Excitatory synaptic transmission onto VTA-projecting LHB neurons of animals with cLH had a significantly lower failure rate than that of wild-type control animals (Fig. 3c;

Figure 4 | DBS in the LHB suppresses excitatory synaptic transmission and reverses learned helplessness. **a**, Left, example excitatory postsynaptic potentials (EPSPs) recorded from a VTA-projecting LHB neuron before (1), during (2) and after (3) stimulation mimicking DBS. Arrows indicate when paired pulses were given. Right, mean EPSP slope at indicated time points before, during and after DBS: (1) before (first pulse 1.1 ± 0.1 , second pulse 0.9 ± 0.1 , $n = 6$ (4 animals)); (2) during (first pulse 0.15 ± 0.07 , second pulse 0.03 ± 0.03 , $P < 0.001$ for both pulses compared with those in 1, Student's t -test); and (3) after (first pulse 1.2 ± 0.3 , second pulse 0.9 ± 0.2).

b, A schematic diagram showing the experimental procedures. **c**, Number of lever presses (left) and test completion time (right) for animals that received DBS or sham stimulation in the LHB or DBS in the LPLR (lateral post-thalamic nuclei, lateroventral), before (baseline) or after DBS of different intensities. For DBS in the LHB ($n = 9$), lever press baseline 1.2 ± 0.4 ; 150 μ A session 3.9 ± 1 ; and 300 μ A session ($n = 8$) 5.8 ± 2 . For DBS in the LHB ($n = 9$), test completion time baseline 19.9 ± 0.4 ; 150 μ A session 17.1 ± 1 ; 300 μ A session ($n = 8$) 15.4 ± 2 . For sham stimulation in the LHB ($n = 14$), lever press baseline 1.8 ± 0.5 ; 150 μ A session 2.4 ± 0.9 ; and 300 μ A session 2.4 ± 1 . For sham stimulation in the LHB, test completion time baseline 19.2 ± 0.5 ; 150 μ A session 18.7 ± 0.9 ; and 300 μ A session 18.5 ± 1.1 . For DBS in the LPLR ($n = 7$), lever press baseline 0.4 ± 0.2 ; 150 μ A session 2 ± 1 ; and 300 μ A session 1.1 ± 0.3 . For DBS in the LPLR, test completion time baseline 20.6 ± 0.2 ; 150 μ A session 19 ± 1.2 ; and 300 μ A session 19.9 ± 0.3 . For DBS in the LHB group, *, $P < 0.05$ compared with baseline (bootstrap). For the sham and DBS in LPLR groups, $P > 0.05$ for both measurements at both sessions compared with baseline (bootstrap). **d**, Immobility during the forced swim test. For DBS, day 1, 0.06 ± 0.01 ; day 2, 0.06 ± 0.01 ; $n = 9$. For sham, day 1, 0.05 ± 0.01 ; day 2, 0.17 ± 0.05 ; $n = 10$. DBS versus sham on day 2, ***, $P < 0.001$ (bootstrap). **e**, Representative cresyl violet staining of coronal brain sections after DBS in the LHB or LPLR. Arrows indicate the electrode track in the LHB (top) or LPLR (bottom). Dashed lines indicate the border of the habenula. Scale bars, 1 mm. **a**, **c**, **d**, Error bars, s.e.m.

$n = 7-9$, $P < 0.001$, bootstrap method). These results indicate that the excitatory synaptic inputs onto VTA-projecting LHB neurons of helpless animals have a higher synaptic release probability; therefore, repeated stimulation can deplete synaptic vesicles faster and more efficiently in helpless animals.

One treatment for clinical depression that is currently under evaluation is deep brain stimulation (DBS). This therapy consists of continuously delivered high-frequency electrical stimulation to various brain regions^{8,18}. In a recent clinical case, DBS of the LHB produced a marked remission of treatment-resistant depression⁹. Notably, depression recurred when DBS was stopped (in two accidental episodes⁹). To examine the cellular effects of DBS, we used brain slices and recorded synaptic transmission onto VTA-projecting LHB neurons, which was evoked by placing a stimulating electrode in the LHB. After a baseline period of evoked transmission, a DBS protocol used in patients (trains consisting of seven stimuli at 130 Hz followed by a 40 ms interval with no stimuli) was continuously delivered through the same stimulation electrode; stimuli were interleaved with the DBS trains, allowing us to monitor evoked synaptic transmission (see Methods for details). The DBS protocol produced a marked depression of excitatory synaptic transmission, which persisted for the DBS protocol period and was reversed on cessation of the DBS protocol (Fig. 4a). Thus, a DBS protocol can effectively reduce excitatory synaptic transmission onto VTA-projecting LHB neurons.

We wished to test whether reducing synaptic drive onto LHB neurons can modulate helpless behaviour. Remarkably, when the DBS protocol that was used in brain slices was delivered to the LHB in animals with aLH, it markedly ameliorated their helpless behaviour, as indicated by an increase in escape behaviour (Fig. 4b, c and Supplementary Fig. 5a). This effect was dependent on both the intensity of stimulation and the placement of the stimulation electrode: stimulating at 300 μ A had a stronger behavioural effect and affected a larger volume within the LHB than stimulating at 150 μ A (Fig. 4b, c and Supplementary Fig. 5b, c). In addition, only if the electrode was placed in the LHB, but not in the nearby thalamus, did DBS reverse the helplessness (Fig. 4b, c, e). Furthermore, DBS in the LHB, but not sham stimulation in the LHB, prevented the increase in immobility in the forced swim test (Fig. 4d). Thus, suppression of synaptic transmission at the LHB through DBS can acutely reverse helpless behaviour in rats.

Several changes in neural function have been identified in depressed humans and rodent models of depression, probably owing to the multifaceted nature of depressive disorders¹⁹⁻²³. The recent identification of the LHB as a brain region in monkeys that can encode disappointment and expectation of negative conditions^{2,3} led us to investigate its role in the learned helplessness rodent model of depression. Our findings indicate that excitatory synaptic activity onto VTA-projecting neurons in the LHB may be a key modulator of learned helplessness. The two learned helplessness models examined showed potentiated excitatory synaptic activity onto these neurons. Interestingly, the major modification was an increase in the proportion of cells showing high-frequency mEPSCs (from 2% to 14–20%) in animals with learned helplessness. This finding suggests that large changes in a small proportion of cells in the LHB may be capable of modifying an animal's behaviour. A crucial role for transmission onto LHB neurons is further supported by the strong correlation between the potentiation of synaptic transmission onto VTA-projecting LHB neurons and an individual animal's helpless behaviour. Given the presynaptic nature of synaptic potentiation, we examined the effects of synaptic depression by repeated afferent stimulation, a protocol that mimics clinically used DBS. Reducing synaptic transmission onto LHB neurons through a DBS protocol led to acute reversal of learned helplessness. Suppression of transmission onto VTA-projecting LHB neurons probably had a role in mediating this beneficial effect, although modulation of LHB neurons, or axons of passage, projecting to other targets may also be involved.

Our study provides cellular mechanisms that may explain previously reported phenomena: the increase in LHB metabolic activity

observed in humans who are depressed^{24,25} and in animal models of depression^{26,27}; and that lesion^{28,29} or pharmacological silencing³⁰ of the LHB can modulate depression-like symptoms in animal models. Our findings suggest an aberrant cellular process that has not previously been examined in the context of mood disorders and that may be crucial in the aetiology of depression. Future studies aimed at determining the changes in molecular signalling that underlie the synaptic hyperactivity onto LHB neurons may lead to novel and effective treatments able to reverse some forms of depressive disorders.

METHODS SUMMARY

Standard surgical procedures were followed for the *in vivo* injection of retrograde tracers. Rats with cLH were bred as described previously^{11,12}. To prepare animals with aLH, rats were exposed to a learned helplessness 'training session' after *in vivo* injection of retrograde tracers into the VTA. This session consisted of inescapable, uncontrollable electric foot shocks, with random shock duration and unpredictable inter-shock intervals. Control animals were placed in the shocking chamber in parallel, without being shocked. To prepare the cLHms group, rats with cLH were treated with brief, escapable foot shock. Acute brain slices from the various groups were prepared for electrophysiological recordings. To evaluate learned helplessness behaviour, we used both a lever-pressing task and an active avoidance task. In the lever-pressing task, a lever was added to the shocking chamber in the testing session. Foot shock was terminated if the animal pressed the lever. The active avoidance task was performed in a shuttle box equipped with an electrical grid floor and a door separating the two halves. Foot shock was terminated if the animal crossed to the other side of the cage. For the forced swim test, the animal was forced to swim in a cylinder of water at 25–26 °C, and the animal's immobility in the water was measured. To test the effect of DBS on learned helplessness behaviour, rats were first trained and tested, and those that met the learned helplessness criteria were chosen for electrode implantation in the LHB. After recovery from surgery, rats underwent a training session followed by a 'baseline' learned helplessness test. DBS at different intensities was subsequently applied, and animals were tested again for the learned helplessness behaviour. To test the effects of DBS on the forced swim test, DBS in the LHB was applied before the second day of swimming to determine its effect on the animal's immobility.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 October 2009; accepted 7 December 2010.

1. Krishnan, V. & Nestler, E. J. The molecular neurobiology of depression. *Nature* **455**, 894–902 (2008).
2. Matsumoto, M. & Hikosaka, O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**, 1111–1115 (2007).
3. Matsumoto, M. & Hikosaka, O. Representation of negative motivational value in the primate lateral habenula. *Nature Neurosci.* **12**, 77–84 (2009).
4. Hikosaka, O. The habenula: from stress evasion to value-based decision-making. *Nature Rev. Neurosci.* **11**, 503–513 (2010).
5. Ji, H. & Shepard, P. D. Lateral habenula stimulation inhibits rat midbrain dopamine neurons through a GABA_A receptor-mediated mechanism. *J. Neurosci.* **27**, 6923–6930 (2007).
6. Nestler, E. J. & Carlezon, W. A. Jr. The mesolimbic dopamine reward circuit in depression. *Biol. Psychiatry* **59**, 1151–1159 (2006).
7. Krishnan, V. *et al.* Molecular adaptations underlying susceptibility and resistance to social defeat in brain reward regions. *Cell* **131**, 391–404 (2007).
8. Mayberg, H. S. Targeted electrode-based modulation of neural circuits for depression. *J. Clin. Invest.* **119**, 717–725 (2009).
9. Sartorius, A. *et al.* Remission of major depression under deep brain stimulation of the lateral habenula in a therapy-refractory patient. *Biol. Psychiatry* **67**, e9–e11 (2010).
10. Maier, S. F. Learned helplessness and animal models of depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **8**, 435–446 (1984).
11. Henn, F. A. & Vollmayr, B. Stress models of depression: forming genetically vulnerable strains. *Neurosci. Biobehav. Rev.* **29**, 799–804 (2005).
12. Schulz, D., Mirrione, M. M. & Henn, F. A. Cognitive aspects of congenital learned helplessness and its reversal by the monoamine oxidase (MAO)-B inhibitor deprenyl. *Neurobiol. Learn. Mem.* **93**, 291–301 (2010).
13. Porsolt, R. D., Le Pichon, M. & Jalfre, M. Depression: a new animal model sensitive to antidepressant treatments. *Nature* **266**, 730–732 (1977).
14. Lecourtier, L. & Kelly, P. H. A conductor hidden in the orchestra? Role of the habenular complex in monoamine transmission and cognition. *Neurosci. Biobehav. Rev.* **31**, 658–672 (2007).
15. Zhou, T. C. *et al.* The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* **61**, 786–800 (2009).
16. Kauffling, J. *et al.* Afferents to the GABAergic tail of the ventral tegmental area in the rat. *J. Comp. Neurol.* **513**, 597–621 (2009).

17. Zucker, R. S. & Regehr, W. G. Short-term synaptic plasticity. *Annu. Rev. Physiol.* **64**, 355–405 (2002).
18. Sartorius, A. & Henn, F. A. Deep brain stimulation of the lateral habenula in treatment resistant major depression. *Med. Hypotheses* **69**, 1305–1308 (2007).
19. Airan, R. D. *et al.* High-speed imaging reveals neurophysiological links to behavior in an animal model of depression. *Science* **317**, 819–823 (2007).
20. Berton, O. & Nestler, E. J. New approaches to antidepressant drug discovery: beyond monoamines. *Nature Rev. Neurosci.* **7**, 137–151 (2006).
21. Pittenger, C. & Duman, R. S. Stress, depression, and neuroplasticity: a convergence of mechanisms. *Neuropsychopharmacology* **33**, 88–109 (2008).
22. Sahay, A. & Hen, R. Adult hippocampal neurogenesis in depression. *Nature Neurosci.* **10**, 1110–1115 (2007).
23. Mill, J. & Petronis, A. Molecular studies of major depressive disorder: the epigenetic perspective. *Mol. Psychiatry* **12**, 799–814 (2007).
24. Morris, J. S. *et al.* Covariation of activity in habenula and dorsal raphe nuclei following tryptophan depletion. *Neuroimage* **10**, 163–172 (1999).
25. Roiser, J. P. *et al.* The effects of tryptophan depletion on neural responses to emotional words in remitted depression. *Biol. Psychiatry* **66**, 441–450 (2009).
26. Shumake, J. & Gonzalez-Lima, F. Brain systems underlying susceptibility to helplessness and depression. *Behav. Cogn. Neurosci. Rev.* **2**, 198–221 (2003).
27. Caldecott-Hazard, S., Mazziotta, J. & Phelps, M. Cerebral correlates of depressed behavior in rats, visualized using ^{14}C –2-deoxyglucose autoradiography. *J. Neurosci.* **8**, 1951–1961 (1988).
28. Amat, J. *et al.* The role of the habenular complex in the elevation of dorsal raphe nucleus serotonin and the changes in the behavioral responses produced by uncontrollable stress. *Brain Res.* **917**, 118–126 (2001).
29. Yang, L. M. *et al.* Lateral habenula lesions improve the behavioral response in depressed rats via increasing the serotonin level in dorsal raphe nucleus. *Behav. Brain Res.* **188**, 84–90 (2008).
30. Winter, C. *et al.* Pharmacological inhibition of the lateral habenula improves depressive-like behavior in an animal model of treatment resistant depression. *Behav. Brain Res.* **216**, 463–465 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Deisseroth for help and suggestions, A. Gifford and A. Biegon for sharing equipment and laboratory space, and members of the Malinow Lab and Li Lab for discussions. This study was supported by the Dana Foundation (B.L.), the Biobehavioral Research Awards for Innovative New Scientists (BRAINS) from the National Institute of Mental Health, National Institutes of Health (1R01MH091903-01) (B.L.) and the Shiley-Marcos Endowment (R.M.).

Author Contributions B.L., J.P., M.M. and C.C. contributed equally to the study. B.L., J.P., M.M., C.C., C.D.P. and D.S. performed and analysed the experiments. C.C. and B.L. made the figures. B.L., F.H. and R.M. designed the study. B.L. and R.M. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.L. (bli@cshl.edu) or R.M. (rmalinow@ucsd.edu).

METHODS

Animals. Wild-type, male Sprague Dawley rats were purchased from Taconic Farms and allowed to acclimatize to the animal facility for 1–2 weeks before experiments were carried out. The cLH rats were bred as described^{11,12}. The rats were housed under a 12-h light–dark cycle (7 a.m. to 7 p.m. light), with food and water freely available. All procedures involving animals were approved by the Institute Animal Care and Use Committees of Cold Spring Harbor Laboratory, University of California, San Diego, and Brookhaven National Laboratory.

Retrograde labelling of VTA-projecting LHB neurons *in vivo*. Standard surgical procedures were followed for *in vivo* injection³¹. To label the VTA-projecting LHB neurons, we injected *in vivo* Alexa-Fluor-488-conjugated cholera toxin ($2 \mu\text{g} \mu\text{l}^{-1}$; Molecular Probes) or a herpes simplex virus expressing enhanced GFP (HSV–GFP, NeuroVex), both of which are retrograde tracers, into the VTA.

Animals were anaesthetized with isoflurane (Baxter) using an isoflurane vaporizer (Paragon Medical) and positioned in a stereotaxic apparatus that was connected to a computer system with a digital rat brain atlas (Angle Two Stereotaxic System, myNeuroLab.com). Injections of tracer solutions (3–5 injection sites along the vertical axis, 100–200 nl per injection) were delivered with a glass micropipette through a skull window ($2\text{--}3 \text{ mm}^2$) by pressure application (5–12 psi, controlled by a Picospritzer II; General Valve). The injections were performed within the following stereotaxic coordinates: -5.3 mm from bregma; 0.96 mm lateral from the midline; and $8\text{--}8.4 \text{ mm}$ vertical from the cortical surface. Rats were injected subcutaneously with 5 mg kg^{-1} carprofen (an NSAID) after surgery. During procedures, animals were kept on a heating pad and were brought back to their home cages after regaining movement. We waited 2–3 days to allow the retrograde labelling of neurons in the LHB before we killed the animals for experiments.

Preparation of acute brain slices and electrophysiology. Male rats of 40–50 days of age were used for all of the electrophysiology experiments. Animals were anaesthetized with isoflurane, decapitated and their brains quickly removed and chilled in ice-cold dissection buffer (110.0 mM choline chloride, 25.0 mM NaHCO_3 , 1.25 mM NaH_2PO_4 , 2.5 mM KCl, 0.5 mM CaCl_2 , 7.0 mM MgCl_2 , 25.0 mM glucose, 11.6 mM ascorbic acid and 3.1 mM pyruvic acid, gassed with 95% O_2 and 5% CO_2). Sagittal slices ($400 \mu\text{m}$) across the LHB were cut in dissection buffer, by using a VT1000 S vibratome (Leica), and subsequently transferred to a storage chamber containing artificial cerebrospinal fluid (ACSF) (118 mM NaCl, 2.5 mM KCl, 26.2 mM NaHCO_3 , 1 mM NaH_2PO_4 , 20 mM glucose, 4 mM MgCl_2 and 4 mM CaCl_2 , at $22\text{--}25^\circ\text{C}$, pH 7.4, gassed with 95% O_2 and 5% CO_2). After at least 1 h recovery time, slices were transferred to the recording chamber and were constantly perfused with ACSF maintained at 27°C .

Experiments were always performed on interleaved wild-type control and cLH or aLH animals. About three-quarters of the experiments were carried out blinded to the experimental group. These showed the same results as the non-blinded experiments, and the data were combined. Whole-cell patch-clamp recordings were obtained with Axopatch-1D amplifiers (Axon Instruments) onto neurons in the LHB under visual guidance using transmitted light illumination. For evoked EPSCs, synaptic transmission was evoked with a bipolar stimulating electrode placed close to the stria medullaris, typically $>0.2 \text{ mm}$ away from cell bodies. Responses were recorded at holding potentials of -60 mV (for AMPA-receptor-mediated responses) and $+40 \text{ mV}$ (for detection of any NMDA-receptor-mediated responses and measurement of rectification). NMDA-receptor-mediated responses were quantified as the mean current between 110 ms and 160 ms after stimulation. Bathing solution (ACSF) contained 119 mM NaCl, 2.5 mM KCl, 2 mM CaCl_2 , 1 mM MgCl_2 , 26.2 mM NaHCO_3 , 1 mM NaH_2PO_4 , 11 mM glucose, and 0.1 mM picrotoxin, gassed with 5% CO_2 and 95% O_2 , at 27°C (unless otherwise noted). Internal solution for voltage-clamp experiments contained 115 mM caesium methanesulphonate, 20 mM CsCl, 10 mM HEPES, 2.5 mM MgCl_2 , 4 mM Na_2ATP , 0.4 mM Na-GTP, 10 mM Na-phosphocreatine and 0.6 mM EGTA (pH 7.2). Spermine ($100 \mu\text{M}$) was included in the internal solution for measurement of rectification. mEPSCs were recorded at 27°C in the presence of $1 \mu\text{M}$ tetrodotoxin (TTX) and $100 \mu\text{M}$ picrotoxin in sagittal slices and analysed using Mini Analysis Program (Synaptosoft). To isolate miniature inhibitory spontaneous responses (mIPSCs), $1 \mu\text{M}$ TTX, $100 \mu\text{M}$ APV (D-(–)-2-amino-5-phosphonopentanoic acid) and $3 \mu\text{M}$ NBQX were added.

For the experiments in which high-frequency stimulation trains were used to determine presynaptic release probability, QX314 (5 mM) was included in the internal solution to prevent the generation of sodium spikes. To recruit the maximal number of axon terminals that can be stimulated by the high-frequency trains, thereby minimizing the effects of axonal failures and reducing the variability in responses, a low concentration (100 nM) of NBQX was included in the bath ACSF. This allowed stimulation at a higher intensity without evoking large EPSCs that could activate voltage-dependent conductances. For experiments testing the effects of DBS on synaptic transmission onto VTA-projecting LHB neurons, evoked EPSCs were monitored before, during and after a stimulation protocol mimicking

clinical DBS. Stimulation consisted of episodes of 44 trains of stimuli separated by 40 ms. During each train, seven stimuli were applied at a frequency of about 130 Hz. The inter-episode interval was 200 ms, during which two stimuli separated by 50 ms were applied to monitor the amplitude and slope of the EPSP. The DBS protocol and the paired-pulse stimulation were delivered using the same electrode.

Two-photon imaging of dendritic spines. Image acquisition and analysis were described previously^{32,33}. Images were acquired on a custom built dual channel two-photon laser scanning microscope (based on the Olympus FluoView laser scanning microscope) using a Ti:Sapphire Chameleon laser (Coherent), which was mode locked to 910 nm. Full three-dimensional (3D) image stacks were acquired using a $\times 60$ 0.9 NA objective lens at $\times 5$ digital zoom (FluoView software, Olympus), 70 nm per pixel. Each image plane was resampled three times and spaced $0.5 \mu\text{m}$ in the z dimension.

Behavioural paradigms. Methods for the learned helplessness paradigm have been optimized previously³⁴. To prepare animals with learned helplessness (the aLH group) for behavioural testing and electrophysiological recording, rats were exposed to a learned helplessness ‘training session’ 5 days after *in vivo* injection of retrograde tracers into the VTA. This session consisted of 120 inescapable, uncontrollable electric foot shocks at 0.8 mA over 40 min in the shocking chambers (Coulbourn Instruments; chambers were 12 inches wide \times 10 inches deep \times 12 inches high and were controlled by precision adjustable shockers), with random shock duration ranging from 5 s to 15 s and unpredictable inter-shock intervals (ITIs). Experiments were performed on pairs of littermates housed in the same cage. Control animals were placed in the shocking chamber in parallel for 40 min, without being shocked. Electrophysiological recordings on acute brain slices were performed 24–48 h after shocking. Animal identity was coded for blinding the researcher with respect to treatment. To prepare cLH animals exposed to mild stress (the cLHms group), cLH rats were treated with a procedure essentially the same as the active avoidance task (see below), during which animals received an average of 152 ± 27 s of escapable foot shock ($n = 8$), and acute brain slices were prepared after 2 h.

To evaluate learned helplessness behaviour, we used both a lever-pressing task and an active avoidance task. The lever-pressing task was described previously³⁴. Briefly, an illuminated lever was added to the shocking chamber in the testing session, which comprised 15 escapable foot shocks lasting up to 60 s (shorter if terminated by a lever press) over 21 min, and with fixed ITIs of 24 s. The active avoidance task was performed in a shuttle box (20 inches wide \times 10 inches deep \times 12 inches high; Coulbourn Instruments) equipped with an electrical grid floor, a door separating the two halves, and photocell detectors. The shuttle box was placed in a sound-attenuating chamber to minimize external stimuli. Testing was fully automated using Graphic State software (Coulbourn Instruments). Animals were allowed to explore the shuttle box for 5 min, and helpless behaviour was evaluated over 30 trials of unexpected and escapable foot shock (1.2 mA intensity, 10 s duration, with random ITIs of 24 ± 12 s) following a 5-s cue tone. Foot shock was terminated if the animal completely crossed to the other side of the cage. When an animal crossed the cage during the 5-s cue tone presentation, avoidance was scored. If an animal crossed during the 10-s shocks, the mean escape latency was measured. Failure was recorded if no crossing was made during the 10-s shock.

For the forced swim test, animals were forced to swim for 5 min in a cylinder of water (water temperature was $25\text{--}26^\circ\text{C}$; the cylinder was 30 cm in diameter and 40 cm high; the depth of the water was set to prevent animals from touching the bottom with their hind limbs). Animal behaviour was videotaped using a PC6EX3 infrared camera (SuperCircuits). The immobile time each animal spent during the test was manually counted offline, with the evaluator being blind to the treatment of the animals.

Deep brain stimulation (DBS). To prepare animals with learned helplessness for the DBS experiments, animals were first treated with a ‘training’ session and then 24 h later, a ‘testing’ session, as described above. On the basis of the test results, animals that met the criteria (those that pressed the lever only 0–5 times, and took between 16 and 21 min to finish the test) were used to test the effects of DBS. For increased stringency, only lever presses occurring within the first 20 s of shock onset were counted. Fifty two male Sprague Dawley rats were trained and tested, and 26 of these animals met the criteria. Three days later, standard surgical procedures were used to implant bipolar concentric electrodes (8 mm long, 0.8 mm tip diameter; Plastics One) unilaterally into the LHB (coordinates -3.7 mm AP , $\pm 0.7 \text{ mm ML}$ and -5.4 mm DV) in rats that met the criteria.

After 3–5 days recovery from surgery, rats underwent a training session followed by a ‘baseline’ learned helplessness test. Three animals, in which the electrodes were implanted into the LHB, were excluded from further study because their performance did not meet the criteria during the baseline test. Immediately following the baseline test, DBS (seven stimulus trains of 130 Hz, separated by 40 ms intervals; $150 \mu\text{A}$ intensity) in the LHB or the thalamus, or no stimulation (sham), was applied for 1 h. Twenty-four hours later, another 1-h session of DBS

or sham stimulation was given immediately before and during the learned helplessness test. DBS intensity was 150 μ A. Another 24 h later, the final 1-h session of DBS (or sham stimulation) was given at a higher intensity (300 μ A), immediately before and during the final learned helplessness test (see Fig. 4b for a schematic diagram showing the experimental procedures). Only animals with the electrode correctly placed in the LHb or thalamus (LPLR) were included for the respective behavioural analysis.

To test the effects of DBS on the forced swim test, electrodes were implanted in the same way as described above, except that in the DBS group two rats had bilateral implants and that in the sham group three rats had bilateral implants. The rest of the animals had unilateral implants (total animals used for the forced swim test: DBS $n = 9$ and sham $n = 10$). Immobility time was recorded during the first 5 min of a 15-min swimming session on day 1. DBS (150 μ A) or sham stimulation was applied for 1 h following the forced swim test on day 1. Twenty-four hours later (day 2), another 1 h of DBS (150 μ A) or sham stimulation was applied, and immobility time was recorded during the 5-min swimming session.

To determine the volume of tissue affected by DBS in the LHb, in a separate set of experiments, animals were perfused with 4% PFA 2 h after the onset of DBS, and brains were processed for immunohistochemistry to examine Fos expression.

Immunohistochemistry. Immunohistochemistry experiments were performed following standard procedures on 50- μ m brain sections fixed with 4% PFA. The

antibodies used were anti-NeuN antibody (Chemicon), anti-EAAC1 antibody (Chemicon), anti-GABA antibody (Sigma), anti-GAD67 antibody (Chemicon) and anti-Fos antibody (Santa Cruz Biotechnology). After finishing the immunohistochemistry process, images were taken using either an LSM 510 confocal microscope (Zeiss; for double labelling with two colours) or a BX41 histology microscope (Olympus; for single labelling with one colour), using $\times 20$ objectives.

Statistics and data presentation. To compare the means of non-normally distributed data sets, we used a bootstrap procedure. Two data sets (N and M of size n and m) were randomly sampled n and m times, respectively, allowing resampling, and means (N_i and M_i) were generated. This procedure was repeated 10,000 times. If N_j was more than M_j fewer than 5% of the times, then the probability that N is more than M was estimated to be less than 0.05. Similar calculations established probabilities less than 0.01. All other statistical tests are indicated when used. All data are presented as mean \pm s.e.m.

31. Rumpel, S., LeDoux, J., Zador, A. & Malinow, R. Postsynaptic receptor trafficking underlying a form of associative learning. *Science* **308**, 83–88 (2005).
32. Kopec, C. D. *et al.* Glutamate receptor exocytosis and spine enlargement during chemically induced long-term potentiation. *J. Neurosci.* **26**, 2000–2009 (2006).
33. Kopec, C. D., Real, E., Kessels, H. W. & Malinow, R. GluR1 links structural and functional plasticity at excitatory synapses. *J. Neurosci.* **27**, 13706–13718 (2007).
34. Vollmayr, B. & Henn, F. A. Learned helplessness in the rat: improvements in validity and reliability. *Brain Res. Brain Res. Protoc.* **8**, 1–7 (2001).

Prion propagation and toxicity *in vivo* occur in two distinct mechanistic phases

Malin K. Sandberg¹, Huda Al-Doujaily¹, Bernadette Sharps¹, Anthony R. Clarke¹ & John Collinge¹

Mammalian prions cause fatal neurodegenerative conditions including Creutzfeldt–Jakob disease in humans and scrapie and bovine spongiform encephalopathy in animals¹. Prion infections are typically associated with remarkably prolonged but highly consistent incubation periods followed by a rapid clinical phase. The relationship between prion propagation, generation of neurotoxic species and clinical onset has remained obscure. Prion incubation periods in experimental animals are known to vary inversely with expression level of cellular prion protein. Here we demonstrate that prion propagation in brain proceeds via two distinct phases: a clinically silent exponential phase not rate-limited by prion protein concentration which rapidly reaches a maximal prion titre, followed by a distinct switch to a plateau phase. The latter determines time to clinical onset in a manner inversely proportional to prion protein concentration. These findings demonstrate an uncoupling of infectivity and toxicity. We suggest that prions themselves are not neurotoxic but catalyse the formation of such species from PrP^C. Production of neurotoxic species is triggered when prion propagation saturates, leading to a switch from autocatalytic production of infectivity (phase 1) to a toxic (phase 2) pathway.

According to the widely accepted protein-only hypothesis², prions consist of multi-chain forms of misfolded host-encoded cellular prion protein (PrP^C), referred to as PrP^{Sc} (ref. 3) and are thought to propagate by an autocatalytic process of seeded fibrilisation^{4–6}. Prion infections are invariably associated with prolonged, clinically silent incubation periods which in humans may exceed 50 years⁷ and yet are followed by an aggressive, often stereotypic, clinical phase which is typically measured in months¹. In inbred laboratory mouse lines, incubation periods for a defined mouse prion strain, although also prolonged, are remarkably reproducible, with onsets in a group of intracerebrally inoculated mice typically spanning only a few days following an incubation period of six months or more. Incubation periods inversely correlate with PrP^C expression levels: very short incubation periods can be obtained in transgenic mice with high levels of PrP^C overexpression⁸, whereas mice with 50% of wild-type expression have greatly prolonged incubation periods⁹. The basis of neurotoxicity in prion neurodegeneration, its relationship to these silent incubation periods and the near synchronicity of clinical onset, remain obscure.

Recently, a general model has sought to incorporate the key phenomena of prion strain diversity and mutation, transmission barriers and the uncoupling of neurotoxicity and infectivity, while remaining within the constraint of requiring only a single polypeptide to constitute all strains of infectious and toxic species⁶. According to this model, these phenomena can be explained in terms of the kinetics of prion propagation, determined by interplay between prion strain type and the tissue/host environment. Neurotoxicity is mediated by a lethal PrP species, PrP^L (refs 10, 11), which is distinct from PrP^{Sc}, but its formation is catalysed by it. The toxic effects occur when the PrP^L concentration passes a local threshold⁶.

Many previous studies, largely published many years ago, have measured prion titres during the incubation period in rodents. These studies used conventional end-point dilution or incubation

interval time bioassay of prion titre, with small numbers of mice or hamsters per assay, and had a significant margin of error. Such classical methods have reported contradictory findings. Some studies have reported prion titres as low or undetectable in the early incubation period, with an exponential rise before clinical onset and the highest titres in the terminal clinical stage^{12–15}. Indeed, if prion propagation involved a high-order exponential, this might explain the rapidity and synchronicity of clinical onset as prion titre rose explosively in the late incubation period⁶. However, other studies, notably in hamsters, have suggested that prion titres may plateau in brain, albeit late in the incubation period^{16,17}. The occurrence of such a plateau in brain has been unexplained.

The recently developed scrapie cell assay (SCA) now allows much more precise bioassay of Rocky Mountain Laboratory (RML) prions than conventional mouse assay¹⁸. In the course of experimental studies to test the general protein-only model, we have used the SCA to produce an accurate and detailed measurement of prion titres throughout the incubation period in mice with a range of PrP^C expression levels and incubation periods and attempted to correlate these with clinical onset.

We inoculated large groups of the following mice intracerebrally with RML mouse prions (30 µl 1% brain homogenate containing 10^{5.8} intracerebral LD₅₀ (dose lethal to 50% of inoculated animals) units and groups of 5–6 mice were killed at multiple defined time points or at onset of clinical disease (Fig. 1a): inbred FVB/N mice (*Prnp*^{+/+}; wild-type PrP^C expression level); *Prnp* null mice¹⁹ (*Prnp*^{0/0}; no PrP^C expression; FVB/N background); hemizygous *Prnp* null mice (*Prnp*^{+/-}; 50% wild-type PrP^C expression; FVB/N background) and Tg20 transgenic mice²⁰ (eightfold wild-type PrP^C expression level; FVB/N background). As neuropathological controls, groups of mice of each genotype were inoculated with uninfected 1% mouse brain homogenate. Prion titre in mouse brain was determined by SCA or scrapie cell assay in end-point format (SCEPA) for low-titre samples¹⁸.

As expected, prions were only briefly detectable in *Prnp*^{0/0} mice (that are unable to propagate prions²¹) and were undetectable after 10 days by SCEPA, demonstrating rapid clearance of the original inoculum and allowing clear interpretation of prion titres in the early incubation period in the other mouse groups (Fig. 1b). Remarkably, prion titres in wild-type FVB/N mice rose rapidly very early in the incubation period, reaching a plateau long before clinical onset. A similar rapid early rise to the same prion titre (of approximately 10⁸ infectious units) was seen in both Tg20 and *Prnp*^{+/-} mice, the principal difference being in the duration of the plateau phase. In the case of the short incubation period Tg20 mice this was highly truncated, whereas in the long incubation period *Prnp*^{+/-} mice this was protracted and accounted for the large majority of the incubation time.

Therefore, we can now describe two distinct phases of prion propagation during the incubation period of RML prion infection in mice. In phase 1, prions propagate exponentially until a clearly defined limit is reached. After this point, a plateau phase (phase 2) continues until the onset of clinical disease. The length of phase 2 is inversely proportional to the PrP^C expression level (Fig. 1c). Interestingly, the rate of

¹MRC Prion Unit and Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK.

prion propagation in phase 1 is not proportional to PrP^C expression level and the prion titre attained in the brain during phases 1 and 2 is independent of PrP^C expression level over the range we have studied (0.5–8-fold wild type).

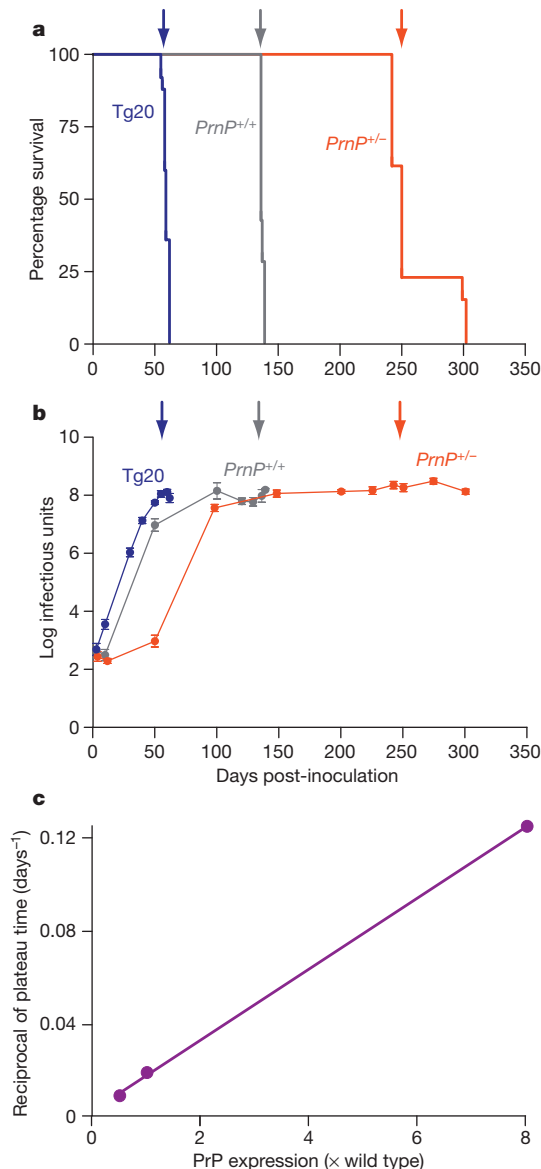


Figure 1 | Survival times and levels of infectivity during prion infection. **a**, Survival curves in mice. Three different lines of mice were infected with RML prions and the survival time measured in days: Tg20 with eightfold overexpression of PrP^C compared to wild type, Prnp^{+/+} with wild-type PrP^C expression level, Prnp^{+/-} with 50% wild-type PrP^C expression. Prnp^{0/0} mice with no PrP^C expression were also inoculated but do not develop disease; data not plotted for clarity, titres were ~2.5 logs at 1 and 3 days and fell to undetectable levels by 10 days. Mean incubation periods, indicated by arrows, were (in days \pm s.d.): Tg20, 59.5 \pm 2.0; Prnp^{+/+}, 137 \pm 1.5; Prnp^{+/-}, 258 \pm 23.5. **b**, Mice were culled at defined time points and prion titre determined. Log tissue culture infectious units are per gram brain; bars indicate s.e.m. and in some cases are smaller than symbol used to designate mean; group sizes were 3–6. Arrows indicate onset of clinical disease. **c**, The reciprocal of the length of the plateau time in days was plotted against the expression level of PrP^C in the mouse lines. Length of phase 2 is inversely proportional to PrP^C expression level. This plot effectively represents the rate at which lethal species of PrP are formed as a function of the concentration of PrP^C. The start of the plateau was defined by the time taken for a linear fit to the logarithmic rise in infectious units at the beginning of each plot to reach the 10⁸ log infectious units plateau level. The end of the plateau was defined by the onset of clinical disease. Plateau lengths were estimated as: Tg20, 8 days; Prnp^{+/+}, 53 days; Prnp^{+/-}, 114 days.

These observations lead to important deductions about prion propagation and neurotoxicity. First, propagation of prion infectivity proceeds in a single phase (phase 1) of exponential autocatalytic conversion. During this process, the addition of PrP^C is only rate-limiting at low expression levels. At high expression levels, that is, for wild-type and the Tg20 overexpressor, the rate of production of infectivity is closely similar. Therefore, the rate-limiting step at high concentrations of PrP^C must be either a structural rearrangement following PrP^C–PrP^{Sc} interaction⁶ or a step in the division of the extended PrP^{Sc}. In phase 2, where there is no increase in infectivity, the rate of formation of the toxic species is directly proportional to PrP^C concentration. In this phase the nature of the conversion process has changed as it must be rate-limited by the addition of PrP^C. That there are closely similar levels of infectivity at the end of phases 1 and 2, irrespective of PrP^C expression level, indicates that there is either a key cofactor, at effectively fixed concentration, or a saturable number of replication sites²², that limits prion production.

Whereas RML prion titres assessed by SCA correlate closely with those obtained by conventional mouse bioassay, these comparisons were made using infectious isolates obtained from terminally clinically affected mice. We considered the possibility that the SCA may be more sensitive in detection of prions generated early in the incubation period, for example if they were more labile and degraded in mouse brain. We therefore performed conventional mouse end-point titration bioassay on samples obtained at 80 days post-inoculation in FVB/N mice (at the start of the plateau) and compared these with those from clinically affected mice (at the end of the plateau). The titre at 80 days was 10^{8.6} LD₅₀ units per ml 10% homogenate, indistinguishable from that at terminal disease.

These studies clearly demonstrate uncoupling of prion infective titre and neurotoxicity, that is, prion titres reach maximal levels in conventional mice with wild-type PrP^C expression long before clinical signs occur. We have proposed previously that neurotoxic species, presumed to be oligomeric forms of PrP (PrP^L), are not on-pathway for prion propagation but produced by a separate but linked pathway in which PrP^{Sc} particles act as a catalytic surface for production of PrP^L (ref. 6). Our current data can be readily accommodated by this general model where production of PrP^L is directly proportional to PrP^C concentration and it is this that principally determines time to onset of clinical disease (Fig. 2). However, alternative models can be envisaged. For example, fission of infectious particles may cease, leading to their further growth but without an increase in their number, or alternatively a vital cellular component may be depleted.

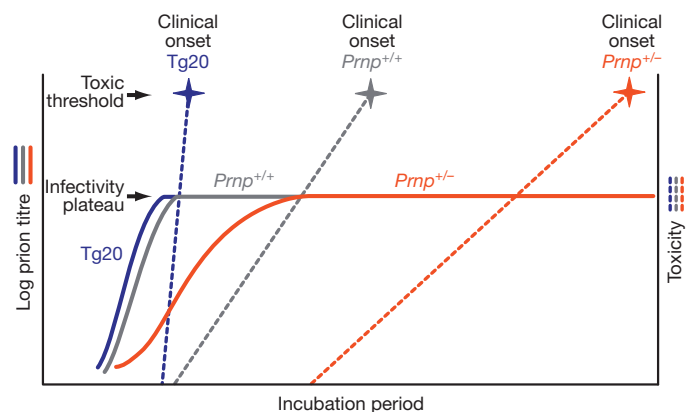


Figure 2 | Kinetics of prion propagation and toxicity. Prion replication (solid lines) is exponential (phase 1) until a limiting titre of prions is attained, which is independent of PrP^C expression level. At this point there is a pathway switch leading to the production of toxic species (dotted lines) at a rate linearly dependent on PrP^C concentration (phase 2). Toxic species do not accumulate until infectivity saturates and clinical signs occur after toxic threshold is crossed.

Whereas disease-related PrP, or PrP^{Sc}, was originally defined in terms of its protease-resistance and detergent insolubility, it is increasingly recognized that there are multiple disease-related forms of PrP, including protease-sensitive species²³, which may indeed comprise the majority of infectivity in some isolates²⁴. As yet, these are poorly defined in physical terms and an agreed nomenclature is lacking. It will be important to determine the relative proportion of these species at different stages of the incubation period. It is possible that a range of PrP species constitute the infectious and toxic entities and these populations may overlap; it is increasingly evident that prion strains constitute a cloud or quasiespecies^{6,25}.

It has also been proposed that PrP^C acts as a receptor for PrP^{Sc} and mediates toxic signalling²⁶. In the light of this, an alternative interpretation could be that variation in PrP^C expression results in a different rate of toxic signalling when prion titres plateau and thereby determines the plateau time before death. However, subclinical prion carrier states exist, where conventional mice with normal neuronal PrP^C expression do not develop clinical signs. These mice live a normal lifespan, despite having prion titres as high as those seen in end-stage clinically sick mice^{10,11,27–29}; such observations are hard to accommodate within a receptor-based model.

METHODS SUMMARY

All procedures were carried out in microbiological containment level 3 facilities. Care of mice was according to institutional guidelines.

RML prion inoculum (I6200) was prepared as a 10% (w/v) brain homogenate in Dulbecco's phosphate buffered saline lacking Ca²⁺ or Mg²⁺ ions (D-PBS) and titred by bioassay in CD1 mice (10^{8.3} intracerebral LD₅₀ per ml)²⁴. I6200 was diluted to 1% (w/v) with D-PBS and 30 µl inoculated intracerebrally into large groups of mice³⁰. Normal CD1 mouse brain homogenate (1% w/v) was inoculated similarly into control mice. Mice used were inbred FVB/N (*Prnp*^{+/+}; wild-type PrP^C expression level), *Prnp* null¹⁹ (backcrossed to FVB/N mice; *Prnp*^{0/0}; no PrP^C expression), hemizygous *Prnp* null FVB/N (*Prnp*^{+/-}; 50% wild-type PrP^C expression) and Tg20 transgenic mice²⁰ (backcrossed to FVB/N mice; eightfold over expression of wild-type PrP^C). Groups of 5–6 mice were killed at multiple defined time points or at onset of clinical disease. Brains were removed and divided sagittally, half-frozen at –70 °C and half-fixed in 10% v/v formal buffered saline. Brain homogenates (10% w/v) were prepared from frozen specimens in D-PBS and stored as aliquots at –70 °C.

Cell culture assays for RML prion infectivity were as described¹⁸. Briefly, PK1 cells were exposed for 3 days in 96-well plates to serial dilutions of 10% (w/v) brain homogenate. A serial dilution of 10% (w/v) RML brain homogenate of known infectivity titre (I6200) was performed in parallel. Subsequently cells were split and passaged appropriately for the scrapie cell assay (SCA) or the scrapie cell assay in end-point format (SCEPA)¹⁸. Infectivity titre of each sample was deduced from the reference preparation. Normal interval regression was used to calculate mean and s.e.m. for two groups that contained a single point below assay sensitivity cut-off. Mouse bioassay was performed by end-point titration using groups of ten FVB/N mice by intracerebral inoculation with 30 µl of serial tenfold dilutions of mouse brain to 10^{–8} (ref. 24); prion titres were calculated using the Reed–Muench formula.

Received 2 September; accepted 17 December 2010.

- Collinge, J. Prion diseases of humans and animals: their causes and molecular basis. *Annu. Rev. Neurosci.* **24**, 519–550 (2001).
- Griffith, J. S. Self-replication and scrapie. *Nature* **215**, 1043–1044 (1967).
- Prusiner, S. B. Novel proteinaceous infectious particles cause scrapie. *Science* **216**, 136–144 (1982).
- Gajdusek, D. C. Transmissible and non-transmissible amyloidosis: autocatalytic post-translational conversion of host precursor proteins to β -pleated sheet configurations. *J. Neuroimmunol.* **20**, 95–110 (1988).

- Come, J. H., Fraser, P. E. & Lansbury, P. T. J. A kinetic model for amyloid formation in the prion diseases: importance of seeding. *Proc. Natl Acad. Sci. USA* **90**, 5959–5963 (1993).
- Collinge, J. & Clarke, A. A general model of prion strains and their pathogenicity. *Science* **318**, 930–936 (2007).
- Collinge, J. et al. Kuru in the 21st century—an acquired human prion disease with very long incubation periods. *Lancet* **367**, 2068–2074 (2006).
- Scott, M. et al. Transgenic mice expressing hamster prion protein produce species-specific scrapie infectivity and amyloid plaques. *Cell* **59**, 847–857 (1989).
- Büeler, H. et al. High prion and PrP^{Sc} levels but delayed onset of disease in scrapie-inoculated mice heterozygous for a disrupted PrP gene. *Mol. Med.* **1**, 19–30 (1994).
- Hill, A. F. et al. Species barrier independent prion replication in apparently resistant species. *Proc. Natl Acad. Sci. USA* **97**, 10248–10253 (2000).
- Hill, A. F. & Collinge, J. Subclinical prion infection. *Trends Microbiol.* **11**, 578–584 (2003).
- Dickinson, A. G., Meikle, V. M. & Fraser, H. Genetical control of the concentration of ME7 scrapie agent in the brain of mice. *J. Comp. Pathol.* **79**, 15–22 (1969).
- Dickinson, A. G. Host-pathogen interactions in scrapie. *Genetics* **79** (Suppl), 387–395 (1975).
- Kimberlin, R. H. & Walker, C. A. Pathogenesis of mouse scrapie: dynamics of agent replication in spleen, spinal cord and brain after infection by different routes. *J. Comp. Pathol.* **89**, 551–562 (1979).
- Kimberlin, R. H. & Walker, C. A. Pathogenesis of scrapie (strain 263K) in hamsters infected intracerebrally, intraperitoneally or intraocularly. *J. Gen. Virol.* **67**, 255–263 (1986).
- Czub, M., Braig, H. R. & Diringer, H. Pathogenesis of scrapie: study of the temporal development of clinical symptoms, of infectivity titres and scrapie-associated fibrils in brains of hamsters infected intraperitoneally. *J. Gen. Virol.* **67**, 2005–2009 (1986).
- Oesch, B. et al. A cellular gene encodes scrapie Prp 27–30 protein. *Cell* **40**, 735–746 (1985).
- Klohn, P., Stoltze, L., Flechsig, E., Enari, M. & Weissmann, C. A quantitative, highly sensitive cell-based infectivity assay for mouse scrapie prions. *Proc. Natl Acad. Sci. USA* **100**, 11666–11671 (2003).
- Büeler, H. et al. Normal development and behaviour of mice lacking the neuronal cell-surface PrP protein. *Nature* **356**, 577–582 (1992).
- Fischer, M. et al. Prion protein (PrP) with amino-proximal deletions restoring susceptibility of PrP knockout mice to scrapie. *EMBO J.* **15**, 1255–1264 (1996).
- Büeler, H. et al. Mice devoid of PrP are resistant to scrapie. *Cell* **73**, 1339–1347 (1993).
- Dickinson, A. G., Fraser, H., Meikle, V. M. & Outram, G. W. Competition between different scrapie agents in mice. *Nat. New Biol.* **237**, 244–245 (1972).
- Safar, J. et al. Eight prion strains PrP^{Sc} molecules with different conformations. *Nature Med.* **4**, 1157–1165 (1998).
- Cronier, S. et al. Detection and characterization of proteinase K-sensitive disease-related prion protein with thermolysin. *Biochem. J.* **416**, 297–305 (2008).
- Li, J., Browning, S., Mahal, S. P., Oelschlegel, A. M. & Weissmann, C. Darwinian evolution of prions in cell culture. *Science* **327**, 869–872 (2010).
- Solforosi, L. et al. Cross-linking cellular prion protein triggers neuronal apoptosis in vivo. *Science* **303**, 1514–1516 (2004).
- Race, R., Raines, A., Raymond, G. J., Caughey, B. & Chesebro, B. Long-term subclinical carrier state precedes scrapie replication and adaptation in a resistant species: analogies to bovine spongiform encephalopathy and variant Creutzfeldt-Jakob disease in humans. *J. Virol.* **75**, 10106–10112 (2001).
- Thackray, A. M., Klein, M. A., Aguzzi, A. & Bujdosó, R. Chronic subclinical prion disease induced by low-dose inoculum. *J. Virol.* **76**, 2510–2517 (2002).
- Thackray, A. M., Klein, M. A. & Bujdosó, R. Subclinical prion disease induced by oral inoculation. *J. Virol.* **77**, 7991–7998 (2003).
- Asante, E. et al. BSE prions propagate as either variant CJD-like or sporadic CJD-like prion strains in transgenic mice expressing human prion protein. *EMBO J.* **21**, 6358–6366 (2002).

Acknowledgements This work was funded by the UK Medical Research Council. We thank J. Wadsworth for critical review of the manuscript, S. Walker for statistical advice, staff of our Biological Services Facility for animal care and technical assistance and R. Young for preparation of figures.

Author Contributions M.K.S., H.A.-D. and B.S. performed the work. M.K.S., A.R.C. and J.C. designed the study and analysed the data. J.C. drafted the manuscript. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.C. (j.collinge@prion.ucl.ac.uk).

Programming the magnitude and persistence of antibody responses with innate immunity

Sudhir Pai Kasturi^{1,2}, Ioanna Skountzou^{1,3}, Randy A. Albrecht⁴, Dimitrios Koutsouanos³, Tang Hua^{1,2}, Helder I. Nakaya^{1,2}, Rajesh Ravindran^{1,2}, Shelley Stewart⁵, Munir Alam⁵, Marcin Kwissa^{1,2}, Francois Villinger^{1,2,6}, Niren Murthy⁷, John Steel⁴, Joshy Jacob^{1,2,3}, Robert J. Hogan⁸, Adolfo García-Sastre^{4,9,10}, Richard Compans^{1,3} & Bali Pulendran^{1,2,6}

Many successful vaccines induce persistent antibody responses that can last a lifetime. The mechanisms by which they do so remain unclear, but emerging evidence indicates that they activate dendritic cells via Toll-like receptors (TLRs)^{1,2}. For example, the yellow fever vaccine YF-17D, one of the most successful empiric vaccines ever developed³, activates dendritic cells via multiple TLRs to stimulate proinflammatory cytokines^{4,5}. Triggering specific combinations of TLRs in dendritic cells can induce synergistic production of cytokines⁶, which results in enhanced T-cell responses, but its impact on antibody responses remain unknown. Learning the critical parameters of innate immunity that program such antibody responses remains a major challenge in vaccinology. Here we demonstrate that immunization of mice with synthetic nanoparticles containing antigens plus ligands that signal through TLR4 and TLR7 induces synergistic increases in antigen-specific, neutralizing antibodies compared to immunization with nanoparticles containing antigens plus a single TLR ligand. Consistent with this there was enhanced persistence of germinal centres and of plasma-cell responses, which persisted in the lymph nodes for >1.5 years. Surprisingly, there was no enhancement of the early short-lived plasma-cell response relative to that observed with single TLR ligands. Molecular profiling of activated B cells, isolated 7 days after immunization, indicated that there was early programming towards B-cell memory. Antibody responses were dependent on direct triggering of both TLRs on B cells and dendritic cells, as well as on T-cell help. Immunization protected completely against lethal avian and swine influenza virus strains in mice, and induced robust immunity against pandemic H1N1 influenza in rhesus macaques.

We designed a nanoparticle-based vaccine, similar to a virus in size and composition. A biodegradable synthetic polymer, poly(D,L-lactico-glycolic acid) (PLGA)⁷, was used to synthesize ~300-nM-sized nanoparticles containing the TLR ligands MPL (TLR4 ligand), R837 (TLR7 ligand), or both ligands, together with an antigen (Supplementary Fig. 1). Immunization of mice with nanoparticles containing MPL and R837 (PLGA(MPL+R837)) plus antigen induced enhanced antibody and T-cell responses, compared to immunization with soluble antigen plus MPL and R837 (data not shown). Consistent with recent observations^{8,9}, delivery of antigen and TLR ligands in separate nanoparticles induced a stronger antibody response than delivery of both in the same nanoparticle (Supplementary Fig. 2). Initially, cohorts of C57BL/6 mice were immunized with nanoparticles containing chicken ovalbumin (OVA) alone (PLGA(OVA)), or (PLGA(OVA)) together with PLGA(MPL), PLGA(R837), or (PLGA(MPL+R837)). OVA emulsified in alum was used as a control. Immunization with PLGA(MPL) or PLGA(R837) plus nanoparticles containing 50 µg or 10 µg of OVA (Supplementary Figs 3a, 4a), induced enhanced

OVA-specific antibody titres after immunization. Notably, there was a synergistic enhancement of the antibody titres in mice that received PLGA(MPL+R837) (Supplementary Figs 3a, 4a). Secondary immunization with the same immunogen 5 weeks later markedly increased titres in all groups, with the synergy effect with PLGA(MPL+R837) still evident, especially at the lower 10 µg dose (Supplementary Figs 3b, 4b). Thus, in all following experiments we used 10 µg of antigen. In addition to OVA, we also used other antigens including the protective antigen (PA) from *Bacillus anthracis*¹⁰ (Supplementary Fig. 5), and haemagglutinin (HA) from avian influenza H5N1 virus¹¹ (Fig. 1). As observed with OVA, there was a synergistic enhancement in the antigen-specific antibody responses after primary and secondary immunization of mice with PLGA(MPL+R837) plus PLGA(PA) (Supplementary Fig. 5a, b), or PLGA(HA) (Fig. 1a, b).

The avidity of antigen-antibody binding is one index of the quality of the antibody response¹². We used a surface plasmon resonance (SPR) assay to assess avidity. Sera from mice immunized with PLGA(MPL+R837) plus PLGA(HA) gave the highest binding response (Fig. 1c). The slower dissociation and higher association rates indicate that immunization with PLGA(MPL+R837) plus PLGA(HA) induced a more enhanced high-affinity antibody response than that induced by PLGA(MPL) or PLGA(R837) plus PLGA(HA). A similar trend was also observed with PA as an antigen (Supplementary Fig. 5c). Consistent with the effects on enzyme-linked immunosorbent assay (ELISA) titres and avidity, mice immunized with PLGA(MPL+R837) plus PLGA(HA) had the greatest neutralization antibody titres (Fig. 1d).

Next we assessed the mechanism by which PLGA(MPL+R837) induced synergistic responses. PLGA(MPL+R837) enhanced the secretion of proinflammatory cytokines by dendritic cells *in vitro*, compared to PLGA(MPL) or PLGA(R837) (Supplementary Fig. 6a). Further, *in vivo* depletion of dendritic cells in CD11c-DTR transgenic mice¹³, or Langerhans cells using the Langerin-DTR mice¹⁴, resulted in diminished antibody titres (Supplementary Fig. 6b, c). These data demonstrate a critical role for dendritic cells in mediating the antibody response to immunization with PLGA(MPL+R837). Signalling via TLR4 and TLR7 is dependent on the adaptor proteins MyD88 or TRIF (also known as TICAM-1); MPL is reported to signal predominantly via TRIF¹⁵. Both MyD88 and TRIF were required for antibody responses stimulated by PLGA(MPL+R837) plus PLGA(OVA) (Supplementary Fig. 7).

B cells express and respond to TLRs^{16,17}. Thus, we determined whether direct triggering of TLRs on B cells was essential for antibody responses. *In vitro* stimulation of naive splenic B cells with PLGA(MPL+R837) synergistically enhanced B-cell proliferation relative to stimulation with PLGA(MPL) or PLGA(R837) (data not shown). To assess this

¹Emory Vaccine Center, Emory University, Atlanta, Georgia 30329, USA. ²Yerkes National Primate Research Center, Emory University, Atlanta, Georgia 30329, USA. ³Department of Microbiology and Immunology, Emory University, Atlanta, Georgia 30322, USA. ⁴Department of Microbiology, Mount Sinai School of Medicine, New York, New York 10029, USA. ⁵Duke Human Vaccine Institute, Duke University Medical Center, Durham, North Carolina 103020, USA. ⁶Department of Pathology, Emory University School of Medicine, Atlanta, Georgia 30322, USA. ⁷The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, USA. ⁸Department of Anatomy and Radiology, College of Veterinary Medicine, University of Georgia, Athens, Georgia 30602, USA. ⁹Department of Medicine, Division of Infectious Diseases, Mount Sinai School of Medicine, New York, New York 10029, USA. ¹⁰Global Health and Emerging Pathogens Institute, Mount Sinai School of Medicine, New York, New York 10029, USA.

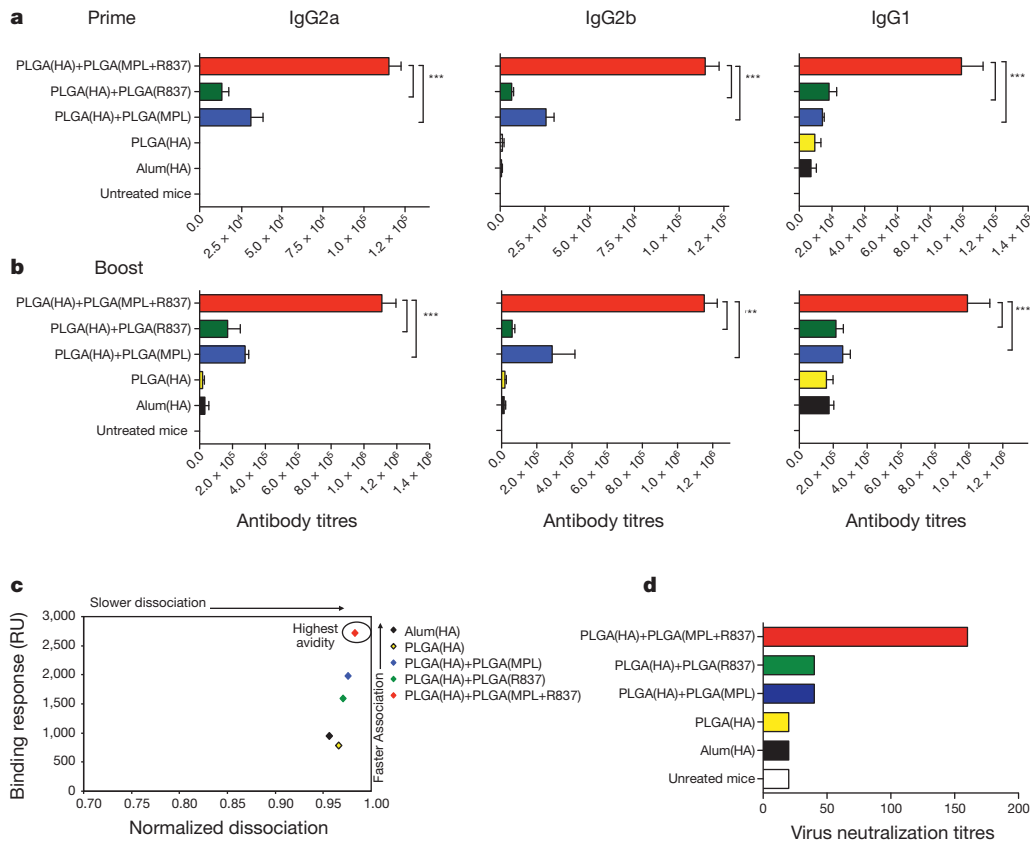


Figure 1 | Combination of MPL and R837 in PLGA nanoparticles mediates synergistic enhancement of antibody responses against H5N1-influenza-derived HA. **a, b**, Antibody titres at 4 weeks after primary and secondary immunization (mean \pm s.e.m. of four independent experiments, with 4–5 mice per treatment group in each experiment) are shown for IgG2a, IgG2b and IgG1 isotypes. *** $P < 0.001$, ** $P < 0.01$ (one-way ANOVA with Bonferroni post-hoc test). **c**, Pooled serum samples from HA-immunized mice at day 28 after

secondary immunization boost immunization were tested for their HA-binding avidity using BIAcore SPR-based protein binding assay. Data are representative of plots from one of two independent experiments. RU, maximal response unit. **d**, Virus neutralization assays were performed with pooled serum samples from treatment groups assayed in duplicates. Results shown are representative titres from one of two independent experiments.

effect *in vivo*, mice lacking B cells (μ MT mice) were reconstituted with B cells from wild-type mice, or *MyD88*^{-/-} or *Trif*^{-/-} mice, and then immunized with PLGA(MPL+R837) plus PLGA(OVA) (Fig. 2a). In μ MT mice reconstituted with wild-type B cells immunization induced a synergistic enhancement of antibody responses (Fig. 2b). However, μ MT mice reconstituted with *MyD88*^{-/-} or *Trif*^{-/-} B cells had diminished antibody titres (Fig. 2c), demonstrating that direct TLR triggering on B cells was required for stimulation of antibody responses. We determined then whether the synergy was dependent on co-expression of TLR4 and TLR7 on the same B cell, or whether there could be complementation between B cells lacking different TLRs. We thus transferred B cells from *Tlr4*^{-/-} mice, or *Tlr7*^{-/-} mice, or a 1:1 mixture of B cells from *Tlr4*^{-/-} and *Tlr7*^{-/-} mice into μ MT mice. Immunization with PLGA(MPL+R837) and PLGA(OVA) demonstrated a requirement for co-expression of both TLRs on the same B cell (Fig. 2d).

Lastly, TLR activation of dendritic cells is known to stimulate antigen-specific CD4⁺ T-helper cells, which are essential for induction of antibody responses¹⁸. Depletion of CD4⁺ T-helper cells before immunization diminished antibody responses (Supplementary Fig. 8). Consistent with this, immunization with PLGA(MPL+R837) and PLGA(OVA) resulted in an enhanced antigen-specific CD4⁺ T-cell response (Fig. 2e). Interestingly, we observed (Fig. 2f) that OVA-specific CD4⁺ T-helper cell responses were substantially reduced in μ MT mice transferred with *MyD88*^{-/-} or *Trif*^{-/-} B cells (in which the antibody responses were severely compromised; Fig. 2c), indicating a requirement for activated B cells in mediating enhanced activation of CD4⁺ T-helper cells, as demonstrated previously¹⁹.

Antibody responses develop along two anatomically and functionally distinct pathways²⁰. The extrafollicular pathway rapidly generates short-lived, antibody-producing cells (plasma cells), and the germinal-centre pathway generates memory B cells and long-lived plasma cells that secrete high-affinity antibody²⁰. We determined whether immunization with nanoparticles containing different adjuvants differentially regulated the two pathways. Thus, mice immunized with the different adjuvants were killed at day 7, lymph nodes isolated and the presence of antibody-producing plasma cells evaluated by immunohistology. There was no apparent difference in the IgG⁺ plasma cells, at day 7, between mice immunized with various adjuvants (Supplementary Fig. 9). We investigated also the kinetics of germinal-centre formation following immunization. Notably, mice immunized with both TLR ligands had a greatly enhanced and sustained germinal-centre response compared to those immunized with a single TLR ligand, (Fig. 3a, b and Supplementary Fig. 9). At day 28 there were approximately 10–12 germinal centres per lymph node and at day 42 there were about 6 germinal centres per lymph node, significantly higher than the numbers observed in lymph nodes of mice immunized with single TLR ligands (Fig. 3a, b). By 8 weeks after immunization the numbers of germinal centres were still substantially higher in the PLGA(MPL+R837) group than in the other groups (Fig. 3b). This demonstrates that PLGA(MPL+R837) preferentially enhances the germinal-centre pathway.

Further, assessment of antigen-specific IgG-secreting plasma cells by the enzyme-linked immunosorbent spot (ELISPOT) assay, indicated no differences in the numbers of plasma cells at day 7 in the

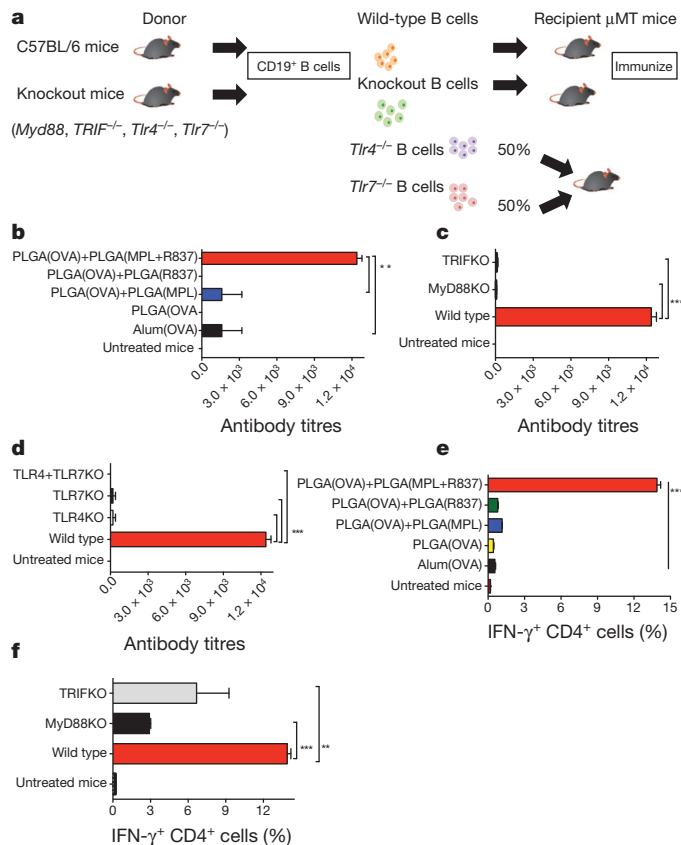


Figure 2 | Synergistic enhancement of antibody responses is dependent on the presence of TLRs on B cells. **a**, B-cell-deficient mice (μ MT mice) were reconstituted with 40×10^6 B cells from C57BL/6 mice or from *MyD88*^{-/-}, *Trif*^{-/-}, *Tlr4*^{-/-} or *Tlr7*^{-/-} mice, or equal numbers of *Tlr4*^{-/-} and *Tlr7*^{-/-} deficient cells to determine whether expression of TLRs and co-expression of TLR4 and TLR7 on the same B cell was necessary for enhancement of antibody responses. Mice were immunized with 10 μ g of OVA encapsulated in PLGA nanoparticles and adjuvants. **b–d**, Mice were bled at day 28 after primary immunization and OVA-specific total IgG antibody responses were determined using ELISA. Antibody titres are shown (mean \pm s.e.m. of two independent experiments, with 3 mice per treatment group in each experiment). KO, knockout. *** $P < 0.001$, ** $P < 0.01$ (one-way ANOVA with Bonferroni post-hoc test). **e**, The magnitude of OVA-specific IFN- γ -producing memory CD4⁺ T cells in the draining lymph nodes of μ MT mice is shown with representative FACS plots. **f**, The magnitude of OVA-specific IFN- γ -producing memory CD4⁺ T cells in the draining lymph nodes is dependent on MyD88 and TRIF expression on B cells. Graphs represent mean frequencies \pm s.d. of triplicate cultures of pooled lymph node cells from one out of two independent experiments.

MPL alone versus MPL+R837 groups (Fig. 3c). At day 28 there were enhanced numbers of plasma cells in the PLGA(MPL+R837) group relative to single TLR-ligand groups. Secondary immunization with the same antigen plus adjuvant resulted in a profoundly enhanced and sustained memory response in the PLGA(MPL+R837) group (Fig. 3c). Notably, there was a persistent plasma-cell response in the lymph node for >1.5 years. Interestingly, there was no corresponding enhancement in the numbers of antigen-specific plasma cells in the bone marrow—a known destination for plasma cells²¹—in the PLGA(MPL+R837) group relative to the single TLR-ligand groups (data not shown). This demonstrates that immunization with PLGA(MPL+R837) preferentially enhances memory B-cell generation. Consistent with this, fluorescence-activated cells sorting (FACS) analysis²² revealed enhanced numbers of isotype-switched, antigen-specific B cells in the PLGA(MPL+R837) group relative to the PLGA(MPL) or PLGA(R837) groups during the memory phase after a secondary boost, but no such enhancement early during the primary response (Supplementary Fig. 10).

The effects of PLGA(MPL+R837) on the germinal-centre pathway might have occurred via early programming of antigen-specific B cells, or as a result of the continued presence of antigen and/or adjuvant. To determine whether there was any early programming of B cells, we isolated isotype-switched, antigen-experienced B cells by FACS at 7 days after immunization with nanoparticles containing various adjuvants plus OVA, and performed microarray analyses to assess their molecular signatures. Notably, there was a great enrichment for genes normally expressed in memory B cells²³, in the cells isolated from mice immunized with PLGA(MPL+R837) plus PLGA(OVA); in contrast, in cells isolated from mice immunized with either PLGA(MPL) or PLGA(R837) plus PLGA(OVA), there was no such enrichment (Supplementary Fig. 11a–c). Such genes included *Bcl2*, *Bcl11a*, *Tank*, several type-I interferon (IFN)-related genes, *Plcg2* and *Cd38*, which are known to have key roles in memory B-cell formation, and several genes that regulate the survival, proliferation and differentiation of germinal-centre B cells, such as *Il17ra*, *Il18r1*, *Pax5*, *Ifngr2*, *Bcor* and *Irf1* (Supplementary Fig. 11a, b). These data indicate that immunization with antigen and MPL+R837 stimulates early programming to the germinal-centre/memory pathway.

Next we determined whether immunization also enhanced antigen-specific memory T-cell responses. There was a synergistic enhancement of OVA-specific IFN- γ -producing CD4⁺ T-helper cell responses at 8 weeks after secondary immunization (Supplementary Fig. 12a, b), but not at days 7 or 14 after primary immunization (data not shown), indicating a preferential effect on the generation of memory T cells. Similar results were observed in mice immunized with HA and PA antigen, even 1.5 years after immunization (Supplementary Fig. 12c–f). PLGA(MPL+R837) immunization also enhanced the antigen-specific CD8⁺ T-cell responses. Although no synergistic enhancement was observed in the frequencies of IFN- γ producing, OVA-specific CD8⁺ T cells at day 7 after primary immunization, there was an increase at day 7 after secondary immunization (Supplementary Fig. 13). Polyfunctional T cells secreting multiple cytokines such as IFN- γ , TNF- α and IL-2 have been implicated in mediating enhanced protection²⁴. We also observed enhanced numbers of triple (IFN- γ , TNF- α and IL-2) and double cytokine (IFN- γ , IL-2) producing CD8⁺ T cells in mice immunized with PLGA(MPL+R837) (Supplementary Fig. 14a, b). Thus, PLGA(MPL+R837) enhances the magnitude and quality of the antigen-specific memory CD4⁺ and CD8⁺ T cells. To assess the relevance of enhanced B- and T-cell responses for protective immunity, we evaluated efficacy of these vaccines in mediating protection in mice, against the 2009 pandemic H1N1 influenza A virus¹¹ (Supplementary Figs 15, 16) and the H5N1 avian influenza virus^{11,25} (Supplementary Fig. 17). In each case, there was enhanced antigen-specific humoral immunity and survival against lethal infection of mice (Supplementary Figs 15–17).

Finally, we assessed the immunogenicity of PLGA(MPL+R837) in non-human primates. In humans and non-human primates, unlike in mice, TLR7 is selectively expressed on plasmacytoid dendritic cells, and not on myeloid dendritic cells². As multiple dendritic-cell subsets seem to be involved in the stimulation of antibody responses by PLGA(MPL+R837) (Supplementary Fig. 6b, c), we used R848, which signals through both TLR7 and TLR8^{2,26}. Further, in humans, although naive B cells do not express TLRs 4, 7 or 8, activated B cells including plasma cells upregulate and respond to these TLRs²⁷. Current human monovalent vaccines against H1N1 influenza contain 15 μ g of HA (effectively 45 μ g of whole inactivated virus (WIV))²⁸. We therefore used 50 μ g of WIV ($\sim 16 \mu$ g HA), as well as a fivefold lower dose ($\sim 3 \mu$ g HA), with and without adjuvant, to determine whether there was a dose-sparing effect. Four cohorts of four rhesus macaques per cohort were immunized subcutaneously with the indicated doses of WIV, with or without PLGA(MPL+R837) or PLGA(MPL+R848) (Fig. 4a). After a single immunization with 10 μ g or 50 μ g of WIV without any adjuvants, there was no detectable antibody response (Fig. 4b–d). In contrast, adjuvanting with PLGA(MPL+R848)

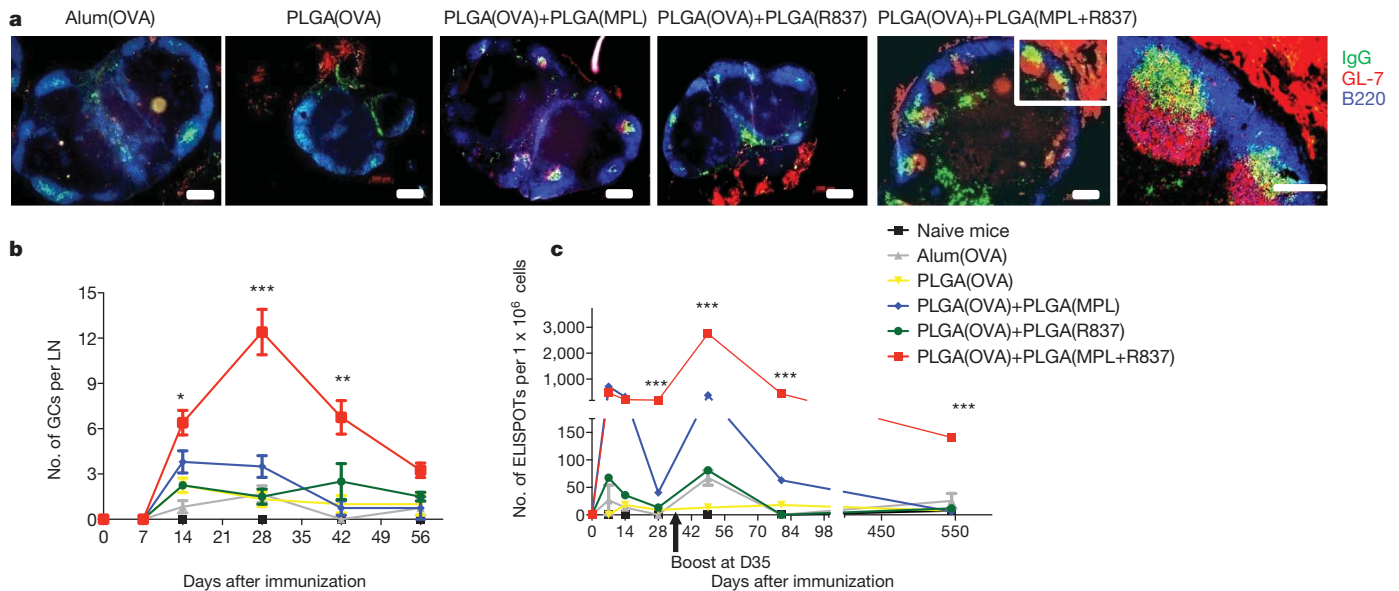


Figure 3 | Immunization with nanoparticles containing MPL and R837 induces persistent germinal centres and long-lived antibody-forming cells in draining lymph nodes. **a**, C57BL/6 mice were immunized with OVA encapsulated in nanoparticles with MPL + R837 plus antigen. Four weeks after primary immunization, draining lymph nodes were excised, tissue sections prepared and stained for germinal centres (GL-7, red; B220, blue; and IgG, green). Images are representative of two independent experiments with draining lymph nodes obtained from 2–3 mice per treatment condition per experiment. Scale bars, 200 μ m for first five panels from left and 100 μ m for

right-hand panel. **b**, Germinal centres (GCs) were counted in lymph node (LN) sections at the time points indicated and represented as mean \pm s.e.m. from 4–6 draining lymph nodes from $n = 2$ –3 mice per treatment group. **c**, ELISPOT assay. Combination of TLR4 and TLR7 ligands has no effect on the short-lived antibody-secreting cells, relative to single TLR ligands, but stimulates long-lived antibody-secreting cells that persist for ~ 1.5 years. Graph represents average spots per 1×10^6 total lymph node cells \pm s.e.m. from duplicate cultures per treatment group. Data are representative of at least 2–3 independent experiments per time point indicated.

induced robust antibody responses, as early as 2 weeks (Fig. 4b–d). Immunization with PLGA(MPL+R837) yielded enhanced binding antibodies, but had more modest effects on HA inhibition and neutralization titres (Fig. 4b–d). Notably, the magnitude of HA inhibition and neutralization titres 28 days after a single immunization with PLGA(MPL+R848) was much greater than 1/40, considered the correlate of protection against influenza in humans (Fig. 4c, d)²⁹. There was at least a fivefold dose-sparing effect, as 10 μ g of antigen plus PLGA(MPL+R848) yielded a much greater response than that

induced by 50 μ g of antigen alone (Fig. 4b–d). Secondary immunizations enhanced the antibody responses in all of the groups, and although PLGA(MPL+R848) still induced the strongest response, PLGA(MPL+R837) also induced responses greater than that required for protection (Fig. 4c, d). Further, the dose-sparing effect was still evident after secondary immunization (Fig. 4b–d).

Here we have described a nanoparticle-based vaccine that resembles a virus in size and composition, and that recapitulates the immunogenicity of live viral vaccines⁵. A notable feature of the immune

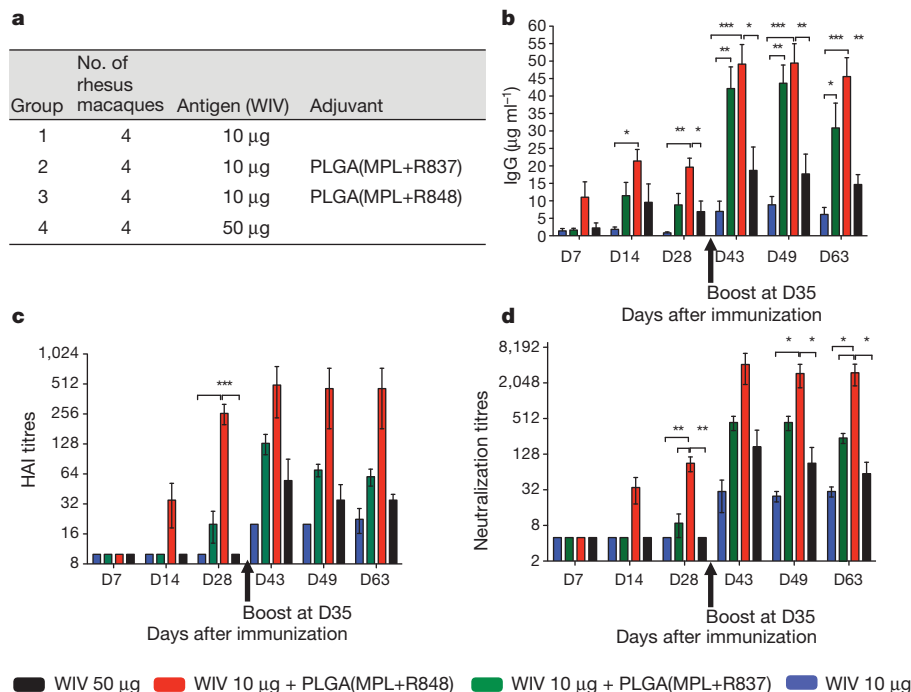


Figure 4 | Immunization of rhesus macaques with 2009 pandemic H1N1 influenza A, whole inactivated virus and nanoparticles containing MPL+R837 or MPL+R848 induces robust humoral immune responses. **a**, Rhesus macaques ($n = 4$) were immunized with 10 μ g of H1N1 WIV with or without nanoparticle-encapsulated MPL+R837 or MPL+R848. Fifty micrograms of MPL and 750 μ g of R837 and R848 encapsulated in nanoparticles were used per macaque. One group of 4 animals was also immunized with 50 μ g of WIV to determine dose-sparing effects mediated by adjuvants. **b**, Antibodies against WIV were analysed as described in Methods and results are represented as mean \pm s.e.m. **c**, HA inhibition (HAI) titres were assayed at indicated time points and are represented as mean \pm s.e.m. **d**, Neutralization titres as represented as the reciprocal of the plasma dilution that decreased the number of plaques formed by the live virus by 50%. Statistical significance was analysed by ANOVA (Bonferroni post-hoc test) and indicated on the figures wherever significant. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$. Data are represented as mean \pm s.e.m.

response stimulated by this vaccine is in the induction of long-lived germinal centres and persistent antigen-specific B- and T-cell responses, similar to that seen with viral infections³⁰. Importantly, although all TLR ligands stimulate primary antibody responses, the combined TLR4 and TLR7 stimulus enhances the germinal-centre pathway of memory B-cell formation and long-lived plasma-cell responses, far more efficiently than stimulation with single TLR ligands (Supplementary Fig. 18). The molecular signatures of antigen-activated B cells isolated early after immunization indicates early programming towards a quasi memory state (Supplementary Fig. 11), but it is also likely that persistent antigen/adjuvants, as evidenced by persistent immune complexes in germinal centres (Fig. 3), have a role. A curious aspect of our data is the requirement for the TLR ligands and antigens to be delivered on two separate particles, consistent with other studies^{8,9}. From a practical perspective, this offers flexibility in coupling a generic adjuvant-containing particle with another particle containing antigen from any pathogen. Further, as each of the components of this vaccine (MPL, R837, PLGA) have been licensed for human use, this vaccine formulation may provide a universal platform for vaccine design against pandemics and emerging infections.

METHODS SUMMARY

Synthesis and characterization of nanoparticles. Antigens were encapsulated in PLGA nanoparticles using a water in oil in water (w/o/w) emulsion technique. MPL was encapsulated in PLGA formulations using an oil in water (o/w) single emulsion process as described before with slight modifications⁸. Sizing of the nanoparticles was conducted using a dynamic light scattering, and protein loading was assessed as described before³¹.

Immunization of mice and non-human primates. Eight-to-twelve week old BALB/c or C57BL/6 mice (Charles River) were immunized with 10 µg of antigen encapsulated in nanoparticles suspended in 200 µl of PBS subcutaneously at the base of the tail. Ten-to-thirteen year old female Rhesus Macaques (7–10 kg) were used and were immunized subcutaneously in the right leg. All procedures were performed in accordance with the Emory School of Medicine Institutional Animal Care and Use Committee guidelines.

Antigen-antibody affinity binding analysis. SPR binding measurements were carried out on a BIAcore 3000 instrument (BIAcore/GE Healthcare).

H5N1 microneutralization assays. This was performed as described previously²⁵.

B-cell ELISPOT. 1×10^6 lymph node cells were serially diluted and cultured overnight in duplicate wells of OVA-coated nitrocellulose-lined 96-well plates (Millipore). Cells were discarded and wells were treated with biotinylated goat anti-mouse total IgG (Southern Biotech) in PBS/0.5% Tween-20 plus 1% FBS for 1.5 h at room temperature (25 °C). Wells were washed and treated with streptavidin alkaline phosphatase (Vector Labs) for another 1.5 h at room temperature. Finally, NBT/BCIP colorimetric substrate for alkaline phosphatase was added to the wells and the reaction was stopped after visualization of purple coloured spots.

Affymetrix GeneChip analysis. Total RNA from sorted B cells was purified using Trizol (Invitrogen). All RNA samples were checked for purity and for integrity and amplified, processed through EXON Module, fragmented and labelled. Labelled targets hybridized to GeneChip Mouse Gene 1.0ST arrays (Affymetrix).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 27 May; accepted 2 December 2010.

- Pulendran, B. & Ahmed, R. Translating innate immunity into immunological memory: implications for vaccine development. *Cell* **124**, 849–863 (2006).
- Kawai, T. & Akira, S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature Immunol.* **11**, 373–384 (2010).
- Pulendran, B. Learning immunology from the yellow fever vaccine: innate immunity to systems vaccinology. *Nature Rev. Immunol.* **9**, 741–747 (2009).
- Querec, T. *et al.* Yellow fever vaccine YF-17D activates multiple dendritic cell subsets via TLR2, 7, 8, and 9 to stimulate polyvalent immunity. *J. Exp. Med.* **203**, 413–424 (2006).
- Querec, T. D. *et al.* Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nature Immunol.* **10**, 116–125 (2009).
- Napolitani, G., Rinaldi, A., Berton, F., Sallusto, F. & Lanzavecchia, A. Selected Toll-like receptor agonist combinations synergistically trigger a T helper type 1-polarizing program in dendritic cells. *Nature Immunol.* **6**, 769–776 (2005).
- Peek, L. J., Middaugh, C. R. & Berkland, C. Nanotechnology in vaccine delivery. *Adv. Drug Deliv. Rev.* **60**, 915–928 (2008).

- Kazzaz, J. *et al.* Encapsulation of the immune potentiators MPL and RC529 in PLG microparticles enhances their potency. *J. Control. Release* **110**, 566–573 (2006).
- Singh, M., Chakrapani, A. & O'Hagan, D. Nanoparticles and microparticles as vaccine-delivery systems. *Expert Rev. Vaccines* **6**, 797–808 (2007).
- Young, J. A. & Collier, J. R. Attacking anthrax. *Sci. Am.* **286**, 48–59 (2002).
- Chen, G. L. & Subbarao, K. Live attenuated vaccines for pandemic influenza. *Curr. Top. Microbiol. Immunol.* **333**, 109–132 (2009).
- Hangartner, L., Zinkernagel, R. M. & Hengartner, H. Antiviral antibody responses: the two extremes of a wide spectrum. *Nature Rev. Immunol.* **6**, 231–243 (2006).
- Jung, S. *et al.* In vivo depletion of CD11c⁺ dendritic cells abrogates priming of CD8⁺ T cells by exogenous cell-associated antigens. *Immunity* **17**, 211–220 (2002).
- Kissenpfennig, A. *et al.* Dynamics and function of Langerhans cells in vivo: dermal dendritic cells colonize lymph node areas distinct from slower migrating Langerhans cells. *Immunity* **22**, 643–654 (2005).
- Mata-Haro, V. *et al.* The vaccine adjuvant monophosphoryl lipid A as a TRIF-biased agonist of TLR4. *Science* **316**, 1628–1632 (2007).
- Bernasconi, N. L., Traggiai, E. & Lanzavecchia, A. Maintenance of serological memory by polyclonal activation of human memory B cells. *Science* **298**, 2199–2202 (2002).
- Pasare, C. & Medzhitov, R. Control of B-cell responses by Toll-like receptors. *Nature* **438**, 364–368 (2005).
- Mitchison, N. A. T-cell–B-cell cooperation. *Nature Rev. Immunol.* **4**, 308–312 (2004).
- van Essen, D., Dullforce, P., Brocker, T. & Gray, D. Cellular interactions involved in Th cell memory. *J. Immunol.* **165**, 3640–3646 (2000).
- McHeyzer-Williams, L. J. & McHeyzer-Williams, M. G. Antigen-specific memory B cell development. *Annu. Rev. Immunol.* **23**, 487–513 (2005).
- Slifka, M. K., Antia, R., Whitmire, J. K. & Ahmed, R. Humoral immunity due to long-lived plasma cells. *Immunity* **8**, 363–372 (1998).
- McHeyzer-Williams, M. G., McLean, M. J., Lalor, P. A. & Nossal, G. J. Antigen-driven B cell differentiation in vivo. *J. Exp. Med.* **178**, 295–307 (1993).
- Luckey, C. J. *et al.* Memory T and memory B cells share a transcriptional program of self-renewal with long-term hematopoietic stem cells. *Proc. Natl Acad. Sci. USA* **103**, 3304–3309 (2006).
- Wille-Reece, U. *et al.* HIV Gag protein conjugated to a Toll-like receptor 7/8 agonist improves the magnitude and quality of Th1 and CD8⁺ T cell responses in nonhuman primates. *Proc. Natl Acad. Sci. USA* **102**, 15190–15194 (2005).
- Steel, J. *et al.* Live attenuated influenza viruses containing NS1 truncations as vaccine candidates against H5N1 highly pathogenic avian influenza. *J. Virol.* **83**, 1742–1753 (2009).
- Jurk, M. *et al.* Human TLR7 or TLR8 independently confer responsiveness to the antiviral compound R-848. *Nature Immunol.* **3**, 499 (2002).
- Dorner, M. *et al.* Plasma cell toll-like receptor (TLR) expression differs from that of B cells, and plasma cell TLR triggering enhances immunoglobulin production. *Immunology* **128**, 573–579 (2009).
- Clark, T. W. *et al.* Trial of 2009 influenza A (H1N1) monovalent MF59-adjuvanted vaccine. *N. Engl. J. Med.* **361**, 2424–2435 (2009).
- Potter, C. W. & Oxford, J. S. Determinants of immunity to influenza infection in man. *Br. Med. Bull.* **35**, 69–75 (1979).
- Bachmann, M. F., Odermatt, B., Hengartner, H. & Zinkernagel, R. M. Induction of long-lived germinal centers associated with persisting antigen after viral infection. *J. Exp. Med.* **183**, 2259–2269 (1996).
- Sah, H. A new strategy to determine the actual protein content of poly(lactide-co-glycolide) microspheres. *J. Pharm. Sci.* **86**, 1315–1318 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Ahmed and B. Rouse for discussion and comments on the manuscript. We thank B. Norris for assistance with FACS analysis, and H. Oluoch for assistance with cryostat sectioning. The work in the laboratory of B.P. was supported by grants U54AI057157, R37AI48638, R01DK057665, U19AI057266, HHSN266200700006C, N01 AI50025 and U19AI090023 from the National Institutes of Health and a grant from the Bill & Melinda Gates Foundation. Work in A.G.-S. laboratories was partly funded by grants HHSN266200700010C, U54AI57158) and U01AI070469 from the National Institutes of Health.

Author Contributions S.P.K. and B.P. designed the study, planned the experiments and analysed the data. B.P. and S.P.K. wrote the manuscript. S.P.K., I.S. and B.P. designed and performed the H1N1 vaccine studies in mice and primates. D.K. assisted with the H1N1 vaccine studies in mice and primates. R.A.A., A.G.-S. and J.S. designed and performed the neutralization assays and challenge experiments with H5N1 vaccine studies in mice. T.H. and R.R. assisted with experiments. H.I.N. performed the microarray analysis. S.S. and M.A. designed and carried out the SPR-based avidity experiments. M.K. assisted with design and execution of mice and non-human primate experiments. N.M. assisted with design of formulations. J.J. assisted with immunohistochemistry and design of experiments. R.J.H. expressed and purified the recombinant H5HA protein. R.C. helped plan and design the H1N1 vaccine study in mice and primates.

Author Information All microarray data are deposited in the Gene Expression Omnibus under accession number GSE25677. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to B.P. (bpulend@emory.edu).

METHODS

Encapsulation of antigens and TLR ligands in nanoparticles. Antigens were encapsulated in PLGA nanoparticles using a water in oil in water (w/o/w) emulsion technique. Briefly, 200 μ l of protein solution (PBS + 0.5% polyvinyl alcohol (PVA) as an excipient (Sigma Aldrich)), OVA grade VI (Sigma Aldrich) at 50 mg ml⁻¹, *Bacillus anthracis* PA (List Laboratories) at 15 mg ml⁻¹, and A/Vietnam/1203/2004 specific haemagglutinin protein (H5HA; affinity chromatography purified from 293 HEK cells) at 15 mg ml⁻¹ were homogenized with 10% w/v PLGA (RG502H, Bohringer Ingelheim) in dichloromethane (200 mg in 2 ml) with the Powergen homogenizer (Fisher Scientific) at speed 5 for 1.5 min. The water in oil emulsion (w/o) was then added to 15 ml of 5% w/v solution of PVA for the second emulsion step identical to the first emulsion process described earlier, at speed 5. The water in oil in water (w/o/w) double emulsion was then subjected to solvent evaporation for 4 h at room temperature (25 °C).

MPL was encapsulated in PLGA formulations using an oil in water (o/w) single emulsion process as described before with slight modifications⁸. MPL (detoxified lipid A, Avanti Lipids) was dissolved in chloroform at 5 mg ml⁻¹ and TLR7 ligand R837 (Invivogen) was dissolved at 10–20 mg ml⁻¹ in DMSO with heating. R848 was dissolved at 12.5 mg ml⁻¹ in dichloromethane. MPL, 0.5 ml at 5 mg ml⁻¹ was added to 200 mg of PLGA polymer dissolved in 2.0 ml of dichloromethane. For particles containing both MPL and R837, 0.5 ml or 5 mg of R837 in DMSO was added to the mixture of PLGA and MPL. For particles containing both MPL and R848, 8 mg of R848 in dichloromethane was added to the mixture of PLGA and MPL. For particles used in non-human primate studies, 0.1 ml of 5 mg ml⁻¹ MPL, 0.5 ml or 10 mg of R837 and 12.5 mg of R848 was added to 200 mg of PLGA polymer in 2.0 ml of dichloromethane. The organic phase containing PLGA with MPL and/or R837 or R848 was homogenized with 15 ml of a 5% w/v PVA solution for 2 min using a speed setting 6 at room temperature. The oil in water emulsion (o/w) was then added to 85 ml of a 5% w/v solution of PVA surfactant to evaporate the organic solvent for 4 h at room temperature. The nanoparticles formed were centrifuged at 3,500g for 20 min and washed with 50 ml of 0.2- μ m filter sterilized, deionized water 3 times. Nanoparticles were snap frozen in liquid nitrogen and lyophilized using a Freezone 2.5L benchtop lyophilizer (Labconco).

Nanoparticle characterization. Sizing of the nanoparticles was conducted using a dynamic light scattering based sizer (90PLUS) from Brookhaven Instruments. Sizes are represented as the mean diameter of the volume average size distribution \pm standard deviation of different batches. Protein encapsulation levels were estimated as described before using a BCA assay (Pierce Biotechnology)³¹. UV-Vis scan using a Cary Win50 UV-Vis spectrophotometer yielded a peak absorbance for R837 at 325 nm (327 nm for R848) and encapsulation was estimated using a standard curve of R837 or R848 in DMSO and NaOH/SDS. MPL encapsulated in nanoparticles was used in splenic dendritic-cell stimulation experiments with known concentrations of soluble MPL yielding identical cytokine production and estimated at 100% encapsulation efficiency.

Mouse dendritic cell culture. CD11c⁺ dendritic cells were isolated from spleens of naive C57BL/6 mice. Spleens were digested with collagenase type IV (Worthington Chemicals), and CD11c⁺ cells were enriched by positive selection using anti mouse CD11c magnetic beads according to manufacturer's instructions (Miltenyi Biotech). Dendritic cells (1 \times 10⁶ cells ml⁻¹) were cultured in 48-well culture plates with PLGA-encapsulated TLR ligands for 24 h.

Mice. *Il6*^{-/-} and *Ticam1*^{1ps2/ps2} strains were obtained from Jackson Labs. *Ifnar*^{-/-} strain was obtained from S. Speck; *MyD88*^{-/-} and *Tlr7*^{-/-} from S. Akira; *Tlr4*^{-/-} from K. Kobayashi.

Immunization of mice and non-human primates. Eight-to-twelve week old BALB/c or C57BL/6 mice (Charles River Laboratories) were immunized with 10 μ g of antigen in nanoparticles (suspended in 200 μ l of PBS) subcutaneously at the base of the tail. TLR ligands were co-delivered with protein-encapsulated nanoparticles. Three milligrams of PLGA particles containing MPL, R837 or MPL+R837 containing 37.5 μ g of MPL and 60 μ g of R837 were used. Imject Alum (Pierce/ThermoScientific) was used to adsorb protein at a 1:1 volume ratio as per the manufacturer's instructions. Mice were bled via the lateral tail vein at regular intervals after primary and secondary immunizations and serum was isolated for analysis of antibody responses by ELISA assays. CD11c⁺ DTR mice¹³ were immunized 24 h after dendritic-cell depletion. Langerin-DTR mice¹⁴ were immunized 3 weeks after depletion, when there is replenishment of Langerin⁺ dendritic cells in the dermis and lymph nodes, but not in the epidermis³². CD4⁺ T cells were depleted with an anti-mouse CD4⁺ antibody (clone GK1.5, provided by R. Mittler) injected at 250 μ g per mouse as described³³. Ten-to-thirteen year old female Rhesus Macaques (7–10 kg) were used. Animals were immunized subcutaneously in the right leg. All animal procedures were performed in accordance with guidelines established by the Emory School of Medicine Institutional Animal Care and Use Committee Guidelines.

Antibody ELISA. Ninety-six-well Nunc maxisorp plates were coated with 100 μ l of 20 μ g ml⁻¹ of OVA, 1 μ g ml⁻¹ of PA or 0.5 μ g of H5HA protein overnight at 4 °C. Plates were washed 3 times with PBS/0.5% Tween-20 using a Biotek auto plate washer and blocked with 200 μ l of 4% non-fat dry milk (Biorad) for 2 h at room temperature. Serum samples from immunized mice at the indicated time points were serially diluted in 0.1% non-fat dry milk in PBS/0.5% Tween-20 and incubated on blocked plates for 2 h at room temperature. Detection antibodies were obtained from Southern Biotech. Wells were washed and incubated with anti-mouse IgG2c–HRP (horseradish peroxidase) conjugate (1:2,000), anti-mouse IgG2b–biotin conjugate (1:2,000), anti-mouse IgG1–HRP conjugate (1:5,000) and streptavidin–HRP conjugate (1:5,000) in PBS/0.5% Tween-20 for 2 h at room temperature. Plates were washed and developed using 100 μ l per well of tetramethylbenzidine (TMB) substrate (BD Biosciences) and stopped using 2N H₂SO₄. Plates were analysed using a BioRad plate reading spectrophotometer at 450 nm with correction at 595 nm. Antibody titres were represented as serum reciprocal dilution yielding a \geq 0.1 optical density (OD) value at 450 nm. Antibody levels (ng ml⁻¹) in H1N1 WIV-immunized mice were assayed as previously described³⁴. To analyse WIV-specific antibody levels in rhesus macaques, the assay was performed as described for mice. A capture rhesus IgG (clone SB108A) was used to establish a standard curve with the rhesus IgG standards (catalogue no. 0135-01; Southern Biotech). 2 μ g ml⁻¹ of WIV was used to coat Nunc maxisorp plates overnight and plates were blocked with 4% non-fat dry milk. Rhesus plasma samples were used at appropriate dilutions for prime and boost and incubated for 2 h at room temperature. Plates were washed 5 \times with PBS/0.5% Tween-20 using an automated plate washer and a goat anti-rhesus HRP was used at 1:4,000 dilution for 1 h at room temperature, washed and developed with TMB substrate. Antibody concentrations were calculated from the IgG standard curves and represented as μ g per ml.

BIAcore assay. SPR binding measurements were carried out on a BIAcore 3000 instrument (BIAcore/GE Healthcare), as described previously³⁵. Serum samples were injected at a 1:50 dilution in PBS for 5 min at a flow rate of 10 μ l min⁻¹. Binding to the negative control rPA surface was subtracted from each sample curve and binding response was measured at 15 and 500 s after the end of the injection. As a measure of antigen-specific antibody-binding avidity, maximal response unit (binding RU) and dissociation rates were measured. Maximal binding RU was measured after subtraction of non-specific signal on the control surface at 15 s post-injection and normalized dissociation was calculated as a ratio of late-to-early binding responses over a dissociation phase of 500 s. Following each injection cycle, chip surfaces were regenerated with a short injection of 25 mM NaOH.

H5N1 microneutralization assays. Serum samples from H5HA-immunized mice were tested for their ability to neutralize a recombinant (6:2) A/PR/8/34 influenza A virus expressing the A/Vietnam/1203/2004 H5HA and N1NA in cell culture *in vitro* as described previously²⁵.

Histology and immunofluorescence. Draining inguinal lymph nodes were isolated and snap frozen in moulds containing OCT medium, dropped into 2-methyl butane cooled with liquid nitrogen. Frozen lymph nodes were sectioned at 5 μ m, fixed in ice-cold acetone for 10 min, air dried and stored at -80 °C. Sections were fluorescently stained with Dylight488-labelled anti-mouse total IgG (Jackson ImmunoResearch), Alexa647-labelled anti-mouse B220 or GL-7 (Ebioscience), followed by Alexa555-conjugated streptavidin (Invitrogen). Fluorescent images were captured using the \times 5 and \times 20 objectives on a Zeiss Axioscope (Carl Zeiss).

B-cell ELISPOT. 1 \times 10⁶ lymph node cells were serially diluted and cultured overnight in duplicate wells of OVA-coated nitrocellulose-lined 96-well MAHA ELISPOT plates (Millipore). Cells were discarded and wells were treated with biotinylated goat anti-mouse total IgG (Southern Biotech) in assay buffer (PBS/0.5% Tween-20 + 1% FBS) for 1.5 h at room temperature. Wells were washed and treated with streptavidin alkaline phosphatase (Vector Labs) at 1:500 for 1.5 h at room temperature. NBT/BCIP colorimetric substrate for alkaline phosphatase was added to the wells and the reaction was stopped after visualization of purple coloured spots. Number of ELISPOTS per well were counted using an ImmunoSpot ELISPOT reader and represented as number of ELISPOTS (antibody-secreting cells) per total 1 \times 10⁶ lymph node cells.

CD4⁺ T-cell assays. 1 \times 10⁶ lymph node cells were cultured in a 200 μ l volume with 100 μ g ml⁻¹ of OVA protein or 5 μ g ml⁻¹ of PA or 5 μ g ml⁻¹ of H5HA in 96-well round-bottomed plates for 4 days. Cells were transferred to anti-CD3 (10 μ g ml⁻¹) and anti-CD28 (2 μ g ml⁻¹) coated flat-bottomed 96-well plates for 6 h in the presence of Golgi Plug (1 μ g ml⁻¹) (BD Biosciences). Cells were stained for surface CD4 using PerCP anti-mouse CD4 (clone RM4-5) for 30 min at 4 °C. Cells were washed 3 times with FACS buffer, fixed and permeabilized using BD Cytofix/Cytoperm and stained for intracellular IFN- γ using APC-conjugated anti-mouse IFN- γ (clone XMGI.2) for 30 min in 1 \times BD perm/wash solution at 4 °C. Cells were washed with perm wash followed by FACS buffer and acquired on a

FACS Caliber cytometer. FACS data were analysed using the Flow Jo software (Tree Star).

CD8⁺ T-cell assays. Primary and memory CD8⁺ T-cell responses were evaluated at day 7 after primary and secondary immunizations. Briefly, peripheral blood mononuclear cells (PBMCs) were enriched using a sucrose density gradient separation (Histopaque; Sigma Aldrich) and cultured with OVA-specific MHC class I restricted peptide at $1 \mu\text{g ml}^{-1}$ (SIINFEKL) for restimulation *ex vivo* in the presence of brefeldin A ($5 \mu\text{g ml}^{-1}$). Stimulated cells were stained for intracellular cytokines using established protocols from BD Biosciences as explained for CD4⁺ T-cell staining experiments above. Cells were stained with PerCP-conjugated anti-mouse CD8 α (Ly-2) (clone 53-6.7), APC-conjugated anti-mouse IFN- γ (clone XMG1.2), FITC-conjugated anti-mouse TNF- α (clone MP6-XT22) and PE-conjugated anti-mouse IL-2 (clone JES6-SH4). Labelled cells were acquired on a FACS Caliber cytometer and FACS data were analysed using Flow JO software (TreeStar).

B-cell multicolour flow cytometry. Antibodies used: PE-labelled anti-mouse IgG (Jackson Immunoresearch); PerCP-labelled anti-mouse CD19 (clone 6D5; Biolegend); E-fluor-405-labelled anti-mouse IgD (Ebioscience); Alexa430 succinimidyl ester for live/dead cell discrimination (Invitrogen); biotin-conjugated anti-mouse Ly77/GL7 (clone GL-7; Ebioscience); APC-labelled anti-mouse CD138 (clone 281-2; BD Biosciences); APC-Cy7-labelled anti-mouse TCR- β (clone H57-597; Biolegend); and APC-Cy7-labelled anti-mouse CD11b (clone M1/70, Biolegend). OVA was labelled with Alexa488 2,3,5,6 TFP ester (Invitrogen) as per the manufacturer's instructions. Unlabelled dye was separated using a 30K cut-off membrane-specified Amicon Ultra4 centrifugal filter (Millipore). Briefly, collagenase-digested lymph nodes from immunized mice were counted for absolute cell numbers. Cells were first stained with Alexa430 succinimidyl ester for dead-cell staining in PBS for 30 min at 4 °C. Cells were washed 2 times with FACS buffer (PBS with 5% FBS) and stained with PE-labelled anti-mouse total IgG for 30 min at 4 °C. Cells were washed 2 times and labelled with the all the above mentioned anti-mouse surface proteins. Qdot655-labelled streptavidin was used to label the biotinylated anti-mouse GL-7 antibodies for another 30 min at 4 °C. Stained cell samples were fixed with BD Cytotfix (BD Biosciences) and acquired on a LSR-II cytometer (BD Biosciences). FACS data were analysed on Flow JO software. For B-cell sorting assays, cells were fluorescently labelled as described above and sorted on a FACS Aria cell sorter (BD Biosciences).

H1N1 virus stock preparation. Madin-Darby canine kidney (MDCK) cells (ATCC CCL 34, American Type Culture Collection) were maintained in DMEM (Mediatech) containing 10% FBS. Stocks of influenza virus strains were prepared by inoculation with H1N1 swine-origin A/California/04/09 strain in allantoic fluid, in 10- or 11-day-old embryonated hen's eggs. Virus stocks were harvested from the allantoic fluid. The purity of the virus was determined by SDS-PAGE in combination with Coomassie blue stain and electron microscopy. The HA activity was determined using chicken red blood cells 0.5% w/v in PBS pH 7.2 as previously described³⁶. The purified virus was inactivated with formaldehyde at a final concentration of 0.01% (v/v), incubated for 72 h at 4 °C, and then dialysed against PBS buffer. Inactivation of virus was confirmed by inoculation of the virus into 10-day-old embryonated hen's eggs and plaque assay in MDCK cells³⁷. We generated mouse adapted A/California/04/09 virus by five serial passages in BALB/c mice. We then determined the LD₅₀ for this virus using the Reed-Muench formula³⁸. For the challenge and infection studies the mice were anaesthetized with isoflurane and then infected with virus by intranasal instillation.

H1N1 HA inhibition assays. Determined as described previously³⁴.

H5N1 and H1N1 lethal challenge infections. Female BALB/c mice immunized with the indicated vaccine schedule were anaesthetized by intraperitoneal injection of ketamine/xylazine and then intranasally infected with 1,000 LD₅₀ of recombinant A/Vietnam/1203/2004 within the enhanced BSL3 Emerging Pathogens Facility at Mount Sinai School of Medicine. To determine survival rates post challenge with mouse adapted A/California/04/09 H1N1 virus, 5 female BALB/c mice per treatment group were challenged 8 weeks after immunization by intranasal instillation of 30 μl of $20 \times \text{LD}_{50}$ live virus after anaesthesia with isoflurane and monitored for morbidity and mortality up to 14 days. Animals were humanely killed and reported as dead if body weight loss achieved 25%. All H5N1 animal procedures were performed in accordance with guidelines established by the Mount Sinai School of Medicine Institutional Animal Care and Use Committee and National Institutes of Health for the care and use of laboratory animals.

H1N1 neutralization assays with primate plasma. Sera from non-human primates immunized with swine-origin H1N1 2009 were serially diluted and mixed with 100 p.f.u. of MDCK-grown homologous virus for 1 h at room temperature. The mixture was further added to an MDCK-cell monolayer and incubated for 45 min at room temperature. The inoculum was removed, wells were overlaid with DMEM agar and incubated for 2 days at 37 °C in a 5% CO₂ humidified incubator. Then plates were fixed with 0.25% glutaraldehyde and stained with 1% crystal violet

in 20% ethanol, and plaques were counted. Neutralizing antibody titres were determined as the reciprocal of the serum dilution that decreased by 50% the number of plaques formed by the live virus.

Affymetrix GeneChip analysis. Total RNA from sorted B cells (1.8×10^4 to 1.2×10^6 cells) was purified using Trizol (Invitrogen). All RNA samples were checked for purity using a ND-1000 spectrophotometer (NanoDrop Technologies) and for integrity by electrophoresis on a 2100 BioAnalyser (Agilent Technologies). The samples were amplified using the Nugen WT Pico Kit (NuGEN Technology) and the target reactions were run with 50 ng of total RNA. The amplification products were processed through the EXON Module (NuGEN Technology), which creates sense-strand cDNA targets. The sense-strand cDNA targets were then fragmented and labelled using NuGEN's FL-Ovation cDNA Biotin Module V2 (NuGEN Technology). Labelled targets were hybridized to GeneChip Mouse Gene 1.0ST arrays (Affymetrix), following standard Nugen protocols for target hybridization to the Affymetrix Gene Arrays. The hybridizations were run for 16 h at 45 °C, 60 r.p.m. in an Affymetrix Hybridization Oven 640. The Cartridge arrays were washed and stained using the Affymetrix Fluidics Stations 450, following Affymetrix protocols. Scanning was performed on an Affymetrix GeneChip 3000 7G scanner, and Affymetrix GCOS software was used to perform image analysis and generate raw intensity data. Two independent sets of samples at day 7 post-treatment were used in our analyses. Each set is comprised of B cells from mice immunized with MPL+R837 or from those immunized with antigen plus MPL or R837 alone. Probe sets of all six samples were normalized by RMA, which includes global background adjustment and quantile normalization. Each set of samples was subsequently normalized by z-score (number of standard deviations from mean) and treated as biological replicates. Affymetrix chip annotation of GeneChip Mouse Gene 1.0ST platform was used to annotate and select probe sets that target a known gene (defined as having an Entrez gene ID). Different probe sets that target the same gene were collapsed by taking the probe set with the highest median expression value across all samples. Student's *t*-test ($P < 0.05$) was used to identify genes differentially expressed in mice immunized with antigen plus MPL+R837 compared to those immunized with antigen plus MPL or R837 alone. A meta-analysis was performed using publicly available microarray data of distinct B-cell subsets (plasma, germinal centre and memory B cells)²³. Purification strategy, RNA-processing method and hybridization strategy can be found in the original publication²³. Raw microarray data (CEL files of samples GSM94747, GSM94762, GSM94763, GSM94764, GSM94765, GSM94766, GSM94767, GSM94768, GSM94769, GSM94771 and GSM94772) were downloaded from the NCBI GEO website (GSE4142) and processed by RMA normalization. Affymetrix chip annotation of the GeneChip Mouse Genome 430 2.0 platform was used to annotate and select probe sets that target a known gene (defined as having an Entrez gene ID). Different probe sets that target the same gene were collapsed by taking the probe set with the highest median expression value across all samples. Student's *t*-test ($P < 0.05$) was used to identify differentially expressed genes (DEGs) between any given two B-cell subsets. Genes were classified as 'plasma DEGs' if they were upregulated in plasma compared to germinal-centre B cells and also upregulated in plasma compared to memory B cells. A similar approach was used to identify 'germinal center DEGs' and 'memory DEGs'.

Genes that were up- or downregulated in mice immunized with antigen plus MPL+R837 compared to those immunized with antigen plus MPL or R837 alone, were cross-referenced to the genes highly expressed in specific B-cell subsets. Fold enrichment was calculated using the formula (common ZX/deg Z)/(subset X/total), where common ZX = number of genes upregulated in treatment Z (combination of MPL+R837 or individual MPL or R837) and also highly expressed in B-cell subset X; deg Z = number of all genes upregulated in treatment Z; subset X = number of all genes highly expressed in B-cell subset X; and total = number of all genes in the chip common to both platforms. In Supplementary Figure 11a, the labelled genes on the right indicate highly expressed genes in the MPL+R837 group that are known to regulate B-cell survival, the germinal-centre response, and differentiation of memory B cells^{39–56}.

32. Henri, S. *et al.* CD207⁺ CD103⁺ dermal dendritic cells cross-present keratinocyte-derived antigens irrespective of the presence of Langerhans cells. *J. Exp. Med.* **207**, 189–206 (2010).
33. den Haan, J. M., Kraal, G. & Bevan, M. J. Cutting edge: lipopolysaccharide induces IL-10-producing regulatory CD4⁺ T cells that suppress the CD8⁺ T cell response. *J. Immunol.* **178**, 5429–5433 (2007).
34. Zhu, Q. *et al.* Immunization by vaccine-coated microneedle arrays protects against lethal influenza virus challenge. *Proc. Natl Acad. Sci. USA* **106**, 7968–7973 (2009).
35. Staats, H. F. *et al.* *In vitro* and *in vivo* characterization of anthrax anti-protective antigen and anti-lethal factor monoclonal antibodies after passive transfer in a mouse lethal toxin challenge model to define correlates of immunity. *Infect. Immun.* **75**, 5443–5452 (2007).
36. Compans, R. W. Hemagglutination-inhibition: rapid assay for neuraminic acid-containing viruses. *J. Virol.* **14**, 1307–1309 (1974).

37. Enioutina, E. Y., Visic, D. & Daynes, R. A. The induction of systemic and mucosal immune responses to antigen-adjuvant compositions administered into the skin: alterations in the migratory properties of dendritic cells appears to be important for stimulating mucosal immunity. *Vaccine* **18**, 2753–2767 (2000).
38. Reed, L. J. & Muench, H. A simple method of estimating fifty percent endpoints. *Am. J. Hyg.* **27**, 493–497 (1938).
39. Badr, G. *et al.* Type I interferon (IFN- α/β) rescues B-lymphocytes from apoptosis via PI3K δ /Akt, Rho-A, NF κ B and Bcl-2/Bcl(XL). *Cell. Immunol.* **263**, 31–40 (2010).
40. Bekeredjian-Ding, I. B. *et al.* Plasmacytoid dendritic cells control TLR7 sensitivity of naive B cells via type I IFN. *J. Immunol.* **174**, 4043–4050 (2005).
41. Thibault, D. L. *et al.* IRF9 and STAT1 are required for IgG autoantibody production and B cell expression of TLR7 in mice. *J. Clin. Invest.* **118**, 1417–1426 (2008).
42. Tovey, M. G., Lallemand, C. & Thyphronitis, G. Adjuvant activity of type I interferons. *Biol. Chem.* **389**, 541–545 (2008).
43. Swanson, C. L. *et al.* Type I IFN enhances follicular B cell contribution to the T cell-independent antibody response. *J. Exp. Med.* **207**, 1485–1500 (2010).
44. Liu, H. *et al.* Functional studies of BCL11A: characterization of the conserved BCL11A-XL splice variant and its interaction with BCL6 in nuclear paraspeckles of germinal center B cells. *Mol. Cancer* **5**, 18 (2006).
45. Smith, K. G. *et al.* *bcl-2* transgene expression inhibits apoptosis in the germinal center and reveals differences in the selection of memory B cells and bone marrow antibody-forming cells. *J. Exp. Med.* **191**, 475–484 (2000).
46. Aiba, Y. *et al.* Preferential localization of IgG memory B cells adjacent to contracted germinal centers. *Proc. Natl Acad. Sci. USA* **107**, 12192–12197 (2010).
47. Zhou, G. & Ono, S. J. Induction of *BCL-6* gene expression by interferon- γ and identification of an IRE in exon I. *Exp. Mol. Pathol.* **78**, 25–35 (2005).
48. Mitsdoerffer, M. *et al.* Proinflammatory T helper type 17 cells are effective B-cell helpers. *Proc. Natl Acad. Sci. USA* **107**, 14292–14297 (2010).
49. Chin, A. I. *et al.* TANK potentiates tumor necrosis factor receptor-associated factor-mediated c-Jun N-terminal kinase/stress-activated protein kinase activation through the germinal center kinase pathway. *Mol. Cell. Biol.* **19**, 6665–6672 (1999).
50. Basso, K. & Dalla-Favera, R. BCL6: master regulator of the germinal center reaction and key oncogene in B cell lymphomagenesis. *Adv. Immunol.* **105**, 193–210 (2010).
51. Kano, G. *et al.* Ikaros dominant negative isoform (Ik6) induces IL-3-independent survival of murine pro-B lymphocytes by activating JAK-STAT and up-regulating Bcl-xL levels. *Leuk. Lymphoma* **49**, 965–973 (2008).
52. Ke, N., Godzik, A. & Reed, J. C. Bcl-B, a novel Bcl-2 family member that differentially binds and regulates Bax and Bak. *J. Biol. Chem.* **276**, 12481–12484 (2001).
53. Airoidi, I. *et al.* Expression and function of IL-12 and IL-18 receptors on human tonsillar B cells. *J. Immunol.* **165**, 6880–6888 (2000).
54. Airoidi, I. *et al.* Heterogeneous expression of interleukin-18 and its receptor in B-cell lymphoproliferative disorders deriving from naive, germinal center, and memory B lymphocytes. *Clin. Cancer Res.* **10**, 144–154 (2004).
55. Hikida, M. *et al.* PLC- γ 2 is essential for formation and maintenance of memory B cells. *J. Exp. Med.* **206**, 681–689 (2009).
56. Nera, K. P. & Lassila, O. Pax5—a critical inhibitor of plasma cell fate. *Scand. J. Immunol.* **64**, 190–199 (2006).

Tumour-infiltrating regulatory T cells stimulate mammary cancer metastasis through RANKL–RANK signalling

Wei Tan^{1†*}, Weizhou Zhang^{1*}, Amy Strasner¹, Sergei Grivennikov¹, Jin Q. Cheng², Robert M. Hoffman^{3,4} & Michael Karin¹

Inflammatory mechanisms influence tumorigenesis and metastatic progression even in cancers whose aetiology does not involve pre-existing inflammation or infection, such as breast and prostate cancers¹. For instance, prostate cancer metastasis is associated with the infiltration of lymphocytes into advanced tumours and the upregulation of two tumour-necrosis-factor family members: receptor activator of nuclear factor- κ B (RANK) ligand (RANKL) and lymphotoxin². But the source of RANKL and its role in metastasis have not been established. RANKL and its receptor RANK control the proliferation of mammary lobuloalveolar cells during pregnancy³ through inhibitor of nuclear factor- κ B (I κ B) kinase- α (IKK- α)⁴, a protein kinase that is needed for the self-renewal of mammary cancer progenitors⁵ and for prostate cancer metastasis². We therefore examined whether RANKL, RANK and IKK- α are also involved in mammary/breast cancer metastasis. Indeed, RANK signalling in mammary carcinoma cells that overexpress the proto-oncogene *ErbB2* (also known as *Neu*)⁶, which is frequently amplified in metastatic human breast cancers^{7,8}, was important for pulmonary metastasis. Metastatic spread of *ErbB2*-transformed carcinoma cells also required CD4⁺CD25⁺ T cells, whose major pro-metastatic function was RANKL production. Most RANKL-producing T cells expressed forkhead box P3 (FOXP3), a transcription factor produced by regulatory T cells, and were located next to smooth muscle actin (SMA)⁺ stromal cells in mouse and human breast cancers. The dependence of pulmonary metastasis on T cells was replaceable by exogenous RANKL, which also stimulated pulmonary metastasis of RANK⁺ human breast cancer cells. These results are consistent with the adverse impact of tumour-infiltrating CD4⁺ or FOXP3⁺ T cells on human breast cancer prognosis^{9,10} and suggest that the targeting of RANKL–RANK can be used in conjunction with the therapeutic elimination of primary breast tumours to prevent recurrent metastatic disease.

RANK signalling controls osteoclastogenesis and bone resorption and has been targeted to prevent bone metastasis in breast and prostate cancers^{11–13}. A humanized anti-RANKL antibody seems to be a more effective inhibitor of bone metastasis than other osteoclast-targeting drugs^{12,13}. To examine whether RANK signalling controls additional aspects of breast cancer metastasis, we used mice on the FVB/N background in which *ErbB2* was expressed under the control of the mouse mammary tumour virus promoter (*MMTV-ErbB2*-transgenic mice)⁶. In these mice, mammary carcinomas are induced by a gene that is overexpressed in 30% of human breast cancers⁷. Because *Rank*-null mice are osteopetrotic and have retarded growth¹¹, we used *MMTV-ErbB2*-transgenic *Rank*^{+/-} heterozygotes on the FVB/N background (*MMTV-ErbB2/Rank*^{+/-}), which were indistinguishable from *MMTV-ErbB2/Rank*^{+/+} mice (on the same background) in appearance, weight

and general health. Although reduced *Rank* gene dosage did not affect primary tumour development, tumoral RANK expression was reduced, and *MMTV-ErbB2/Rank*^{+/-} mice gave rise to 50% fewer lung metastases than did *MMTV-ErbB2/Rank*^{+/+} mice (Fig. 1a). The role of RANK signalling was further examined using an orthotopic tumour model in which primary mammary carcinoma cells from an *MMTV-ErbB2* tumour (PCaM cells) were injected into the number 2 mammary gland of FVB/N mice. Injection of recombinant RANKL, after inoculation with tumour cells, significantly increased the incidence and multiplicity of pulmonary metastases (Fig. 1b).

We also generated an immortalized *ErbB2*-induced mammary carcinoma cell line, MT2, which doubled every 24 h, showed epithelial morphology, and formed colonies in soft agar and mammospheres when grown in suspension without serum (Supplementary Fig. 1a–c). MT2 cells expressed as much RANK as did PCaM cells. However, RANK expression was very low in non-transformed mammary epithelial cells from tumour-free virgins, and neither PCaM nor MT2 cells expressed RANKL (Supplementary Fig. 1d, e). After transplantation to the number 2 mammary gland, MT2 cells underwent pulmonary metastasis that was further enhanced by RANKL injection (Fig. 1c). To ascertain that the results were not affected by a putative ERBB2-elicited immune response, we transplanted MT2 cells to *MMTV-ErbB2* mice on the FVB/N background that had been tolerized to the *ErbB2* oncogene¹⁴. The results were almost identical to those obtained in normal FVB/N mice (Fig. 1c). Silencing of RANK expression (Supplementary Fig. 1f) reduced the pulmonary metastasis of MT2 cells (Fig. 1d) and blocked the response of MT2 cells to RANKL (Supplementary Fig. 2a). Neither administration of RANKL nor silencing of RANK expression significantly influenced the primary tumorigenic growth of MT2 cells, which was identical in FVB/N mice and *MMTV-ErbB2* mice on the FVB/N background (Supplementary Fig. 2b, c). Similar results were obtained by administration of a RANK–Fc fusion protein, which blocks RANK signalling¹⁵, to mice bearing MT2 mammary tumours. Whereas RANK–Fc marginally inhibited tumour growth at the primary site, it greatly reduced pulmonary metastasis (Supplementary Fig. 3).

RANK engagement induced nuclear translocation and phosphorylation of IKK- α molecules that were not associated with IKK- β and IKK- γ (also known as NEMO), the two other subunits of the IKK complex (Supplementary Fig. 4a–c). This finding indicates that free IKK- α or IKK- α that belongs to a complex of a different composition is targeted by RANK. Such RANK-induced IKK- α activation repressed expression of the metastasis inhibitor maspin (also known as SERPINB5) (Supplementary Fig. 5a–c). MT2 cells that overexpressed ectopic maspin (maspin–MT2 cells) showed marginally reduced tumorigenic growth but much less pulmonary metastasis (Supplementary Fig. 5d–g). Treatment with exogenous RANKL stimulated the metastasis of enhanced green fluorescent protein (eGFP)-expressing

¹Laboratory of Gene Regulation and Signal Transduction, Departments of Pharmacology and Pathology, University of California San Diego School of Medicine, 9500 Gilman Drive, Mail Code 0723, La Jolla, California 92093, USA. ²Department of Molecular Oncology, H. Lee Moffitt Cancer Center, Tampa, Florida 33612, USA. ³Department of Surgery, University of California San Diego, San Diego, California 92103-8220, USA. ⁴AntiCancer Inc., San Diego, California 92111, USA. [†]Present address: Pfizer Oncology Research Unit West, San Diego, California 92121, USA.

*These authors contributed equally to this work.

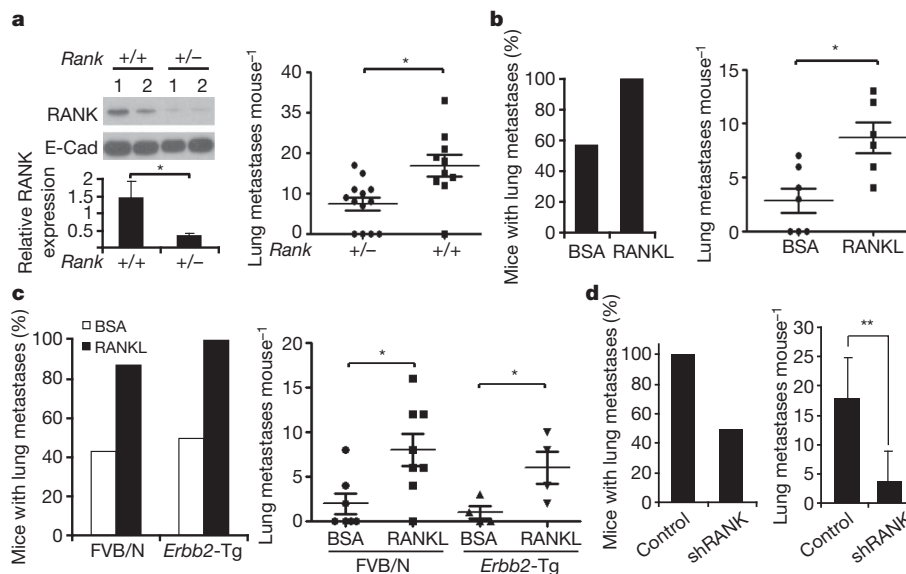


Figure 1 | RANK signalling in mammary carcinoma cells enhances metastasis. **a**, *MMTV-ErbB2/Rank*^{+/+} and *MMTV-ErbB2/Rank*^{+/-} mice were killed 8 weeks after tumour onset. Tumour RANK expression was analysed by immunoblotting (each lane represents a different mouse tumour, top left), quantified by densitometry and normalized to E-cadherin (E-Cad) expression (bottom left; data presented as mean \pm s.e.m.; $n = 6$). Metastasis multiplicity was determined by haematoxylin and eosin (H&E) staining of *MMTV-ErbB2/Rank*^{+/+} ($n = 10$) and *MMTV-ErbB2/Rank*^{+/-} ($n = 13$) mouse lungs. Individual mouse data are shown (right; mean, horizontal line; error bars, s.e.m.). **b**, *MMTV-ErbB2*-transgenic PCaM cells were grafted (1×10^6 cells) into the number 2 mammary glands of FVB/N mice, which were then subjected to twice weekly intratumoural injections of bovine serum albumin (BSA) or recombinant RANKL ($80 \mu\text{g kg}^{-1}$), starting 1 week after inoculation with tumour cells. On day 35, the lungs were isolated, sectioned, stained with H&E

and the metastatic lesions counted. Data are presented as the proportion of mice in each group that developed metastases (left) and as the number of metastases developed by each mouse (right; mean, horizontal line; error bars, s.e.m.; BSA: $n = 7$; RANKL: $n = 6$). **c**, MT2 cells (1×10^6 cells) were grafted as in **b** into FVB/N mice or *MMTV-ErbB2*-transgenic (*ErbB2*-Tg) mice on the FVB/N background, which were then treated with BSA or RANKL, and analysed for pulmonary metastases. Data are presented as incidence of metastases (left) and multiplicity of metastases (right; individual mouse data; mean, horizontal line; error bars, s.e.m.; $n = 7-8$). **d**, Mice were inoculated with MT2 cells (1×10^6 cells) that had been transduced with short hairpin RNA specific for *Rank* (shRANK) or control MT2 cells (1×10^6 cells), and lung metastases were counted 56 days later. Data are presented as in **b** and **c** except that multiplicity is shown as mean \pm s.e.m. ($n = 6$). **a-d**, *, $P < 0.05$; **, $P < 0.01$ (non-parametric Mann-Whitney U test).

MT2 cells but did not affect the growth of primary tumours from either maspin-MT2 or eGFP-expressing MT2 cells; moreover, it did not enhance the weak metastatic activity of maspin-MT2 cells (Supplementary Fig. 5h). As expected⁵, the silencing of IKK- α expression in MT2 cells slowed tumorigenic growth (Supplementary Fig. 6a, b), blocked pulmonary metastasis, even in mice given exogenous RANKL (Supplementary Fig. 6c), and upregulated maspin expression (Supplementary Fig. 6d).

Treatment of primary tumour cells or cells that had reduced *Rank* gene dosage with RANKL did not affect the formation of primary mammospheres (Supplementary Fig. 7a, b). Although RANKL was reported to stimulate metastasis by enhancing cancer cell motility¹⁶, no significant effect of RANKL on motility or invasiveness of MT2 cells was observed (Supplementary Fig. 7c). *In vivo*, treatment with RANKL led to fewer apoptotic cells in late-stage tumours (Supplementary Fig. 7d). Conversely, silencing of IKK- α expression increased the number of apoptotic cells in late-stage tumours by 2.2 fold, and this increase was not reversed by treatment with RANKL (Supplementary Fig. 7e, f). Pre-incubation of tail-vein-injected MT2 cells in RANKL and subsequent *in vivo* treatment with RANKL enhanced the extravasation of MT2 cells into the lungs in a RANK-dependent manner (Supplementary Fig. 7g). *In vitro*, RANKL strongly inhibited the apoptosis of suspended MT2 cells (Supplementary Fig. 7h). Hence, the pro-metastatic effect of RANKL and IKK- α may be due to the increased survival of circulating metastasis-initiating cells.

For *ErbB2*-induced mammary tumours, RANKL expression was detected in the surrounding stroma but not in the carcinoma portion (Fig. 2a and Supplementary Fig. 8a, b) or in metastatic lung nodules, which were surrounded by a few CD5⁺ lymphocytes (Supplementary Fig. 8a). Activated lymphocytes produce RANKL during inflammation¹⁷ and are present in breast tumours¹⁸. We therefore examined

their role in tumour RANKL expression. The staining of sequential tumour sections showed an almost identical distribution of RANKL⁺ cells and CD5⁺ lymphocytes (Fig. 2a). Correspondingly, tumours grown in recombination activating gene 1 (*Rag1*^{-/-}) mice, which lack mature T and B cells (Supplementary Fig. 9), contained less *Rankl* messenger RNA than did tumours in wild-type mice (Fig. 2b) and were devoid of RANKL⁺ cells (Fig. 2a). Decreased *Rankl* mRNA was also seen in tumours grown in *Cd4*^{-/-} mice (Fig. 2c), which lack only CD4⁺ T cells (Supplementary Fig. 9). Tumours in *Cd4*^{-/-} mice were almost completely devoid of RANKL⁺ cells, although they contained a few CD5⁺ cells (Fig. 2a).

To identify further the major RANKL-expressing cell type during metastatic spread, we grew MT2 cells in *Rag1*^{-/-} mammary glands that had been reconstituted with splenic B cells, CD4⁺ T cells or CD8⁺ T cells from FVB/N mice. Reconstitution efficiency was confirmed by flow cytometry (Supplementary Fig. 10). Only CD4⁺ T cells substantially increased tumour *Rankl* mRNA above the low basal level in mock-reconstituted *Rag1*^{-/-} mice (Supplementary Fig. 11a). We analysed B cells, CD4⁺ T cells and CD8⁺ T cells from spontaneous *MMTV-ErbB2* tumours for *Rankl* mRNA and found this to be at the highest level in CD4⁺ T cells (Supplementary Fig. 11b). Likewise, CD4⁺ T cells from MT2 tumours contained much more *Rankl* mRNA than B cells from MT2 tumours (Supplementary Fig. 11c). Despite their lower contribution to *Rankl* mRNA, tumoral CD8⁺ T cells were more numerous than CD4⁺ T cells, and both T-cell subsets were far more abundant than tumoral B cells (Supplementary Fig. 12a). The absence of lymphocytes in *Rag1*^{-/-} mice did not affect the number of tumour-associated macrophages (Supplementary Fig. 12b), and no differences in stromal constituents were found between tumours grown in FVB/N mice and those in *MMTV-ErbB2* mice on the FVB/N background (Supplementary Fig. 2d, e).

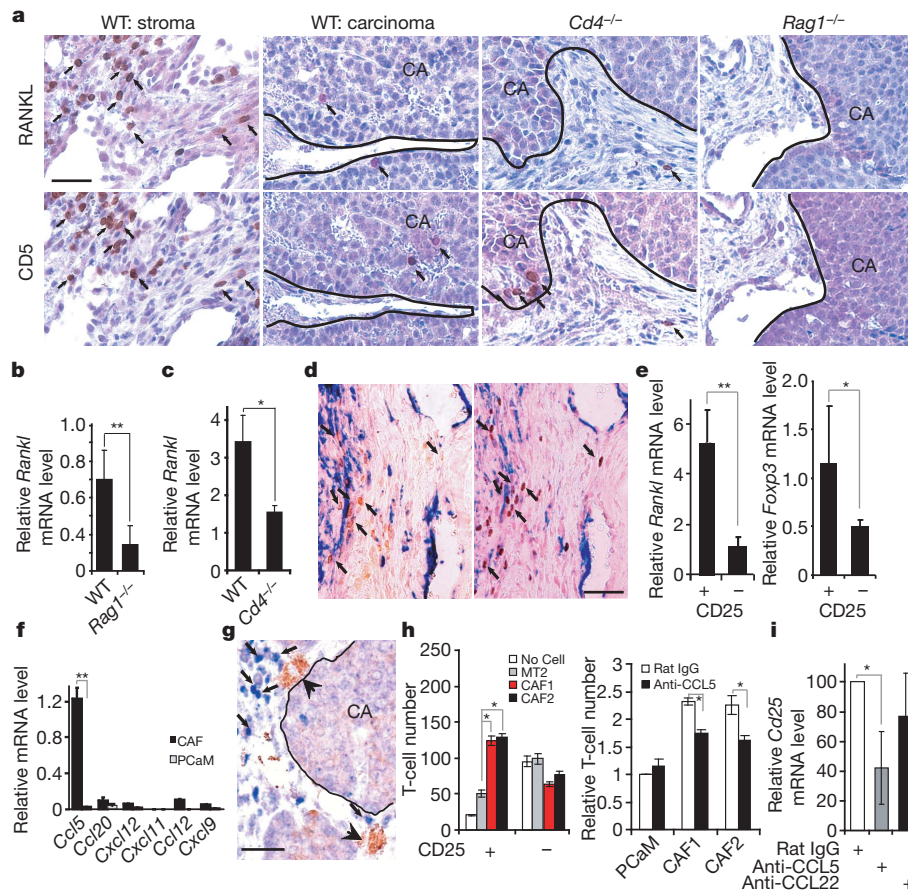


Figure 2 | Expression of RANKL in mammary tumours depends on CD4⁺ T cells. **a**, Mammary glands of indicated mouse strains were inoculated with *MTV-ErbB2*-transgenic PCaM cells (1×10^6 cells). After 56 days, tumours were paraffin embedded. Sequential sections were stained with anti-RANKL and anti-CD5 antibodies and counterstained with haematoxylin. Panels show stroma- and carcinoma-containing (CA) regions separated by a line, when relevant. Arrows indicate CD5⁺ and RANKL⁺ cells. Scale bar, 20 μ m (all images). WT, wild type. **b**, MT2 cells were transplanted to the indicated mouse strains as in **a**. After 56 days, tumours were excised, and *Rankl* mRNA was quantified by quantitative PCR with reverse transcription (RT-PCR) and normalized to cyclophilin mRNA levels. Data are presented as mean \pm s.e.m. ($n = 3$). **c**, *MTV-ErbB2*-transgenic PCaM cells were transplanted to the indicated mouse strains as in **a**. After 56 days, tumours were excised, and *Rankl* mRNA was quantified as in **b**. Data are presented as mean \pm s.e.m. ($n = 3$). **d**, Tumours raised in WT mice as in **a** were excised, and sequential sections were stained with anti-FOXP3 (brown), anti-RANKL (brown) and anti-SMA (blue) antibodies. Arrows indicate RANKL⁺ cells (left) and FOXP3⁺ cells (right). **e**, Tumour-infiltrating CD4⁺CD25⁺ and CD4⁺CD25⁻ T cells were purified and analysed for *Rankl* and *Foxp3* mRNA expression as in **b**. Data are

Among CD4⁺ T-cell types, FOXP3⁺ regulatory T (T_{reg}) cells showed the most similar distribution to RANKL⁺ cells in mammary tumours (Fig. 2d). Notably, T_{reg} cells in breast tumours are associated with an invasive phenotype and poor prognosis^{9,10}. Most FOXP3⁺ and RANKL⁺ cells were in contact with stromal SMA⁺ cells (Fig. 2d). Congruently, CD4⁺CD25⁺ tumoral T cells, which are FOXP3 enriched¹⁹, expressed fourfold more *Rankl* mRNA than CD4⁺CD25⁻ T cells (Fig. 2e). Furthermore, FOXP3 and RANKL were co-localized in tumours (Supplementary Fig. 13a). These results suggest that tumour-infiltrating CD4⁺FOXP3⁺ T_{reg} cells are the most critical cells for maintaining RANKL expression in the micro-environment of metastatic mammary tumours.

The proximity of RANKL⁺ T cells to SMA⁺ cells, which expressed fibroblast markers (Supplementary Fig. 14), prompted us to examine whether SMA⁺ cells express T-cell-attracting chemokines. CC-chemokine ligand 5 (CCL5; also known as RANTES), a T-cell-attracting

presented as mean \pm s.e.m. ($n = 3$). **f**, CAFs and PCaM cells were purified from *MTV-ErbB2*-transgenic tumours. Chemokine mRNAs were quantified as in **b**. Data are presented as mean \pm s.e.m. *Cxcl*, CXC-chemokine ligand.

g, *MTV-ErbB2*-transgenic tumours were sectioned and stained with anti-CCL5 (brown, large arrowheads) and anti-FOXP3 (blue, small arrowheads) antibodies. Scale bar, 20 μ m. **h**, The indicated cell types were plated onto multiwell plates at 2×10^5 cells per well with no addition (left) 3 μ g ml⁻¹ rat IgG (right) or anti-CCL5 antibody (right). Splenocytes from tumour-bearing *MTV-ErbB2*-transgenic mice were added to the upper compartments of the Boyden chambers. After 24 h, the CD4⁺CD25⁺ and CD4⁺CD25⁻ T cells in the bottom compartments were quantified by flow cytometry. CAF1 and CAF2 are two independent preparations. Relative T-cell number is relative to that for PCaM cells treated with IgG. Data are presented as mean \pm s.e.m. ($n = 3$). **i**, Rat IgG, anti-CCL5 or anti-CCL22 antibodies (2 mg kg⁻¹) were injected intraperitoneally into FVB/N female mice bearing MT2 tumours twice weekly. After 1 week, tumours were excised, and *Cd25* mRNA was quantified by quantitative RT-PCR. Data are presented relative to the *Cd25* mRNA levels in rat IgG-treated mice; mean \pm s.e.m. ($n = 4$). **b**, **c**, **e**, **f**, **h**, **i**, *, $P < 0.05$; **, $P < 0.01$ (non-parametric Mann-Whitney *U* test).

chemokine involved in breast cancer metastasis^{20,21}, was expressed by cancer-associated myofibroblasts (CAFs) but not by PCaM cells (Fig. 2f). Expression of CC-chemokine receptor 1 (CCR1), one of the cognate receptors for CCL5 (ref. 22), was increased in tumour CD4⁺ T cells relative to splenic CD4⁺ T cells (Supplementary Fig. 13b, left). Notably, tumoral CD4⁺CD25⁺ T cells expressed more *Ccr1* mRNA than their CD4⁺CD25⁻ T-cell counterparts (Supplementary Fig. 13b, right). Furthermore, tumoral FOXP3⁺ T cells were in close proximity to CCL5⁺ stromal cells (Fig. 2g). Similarly, in *in vitro* experiments, CAFs attracted more CD4⁺CD25⁺ T cells than CD4⁺CD25⁻ T cells from spleens of tumour-bearing mice, although the latter were sixfold more abundant (Fig. 2h, left). Recruitment was partly dependent on CCL5 (Fig. 2h, right). Neutralization of CCL5 also inhibited the recruitment of tumoral CD25⁺ T cells *in vivo* (Fig. 2i).

To examine the requirement for T cells for pulmonary metastasis, we transplanted freshly isolated PCaM cells to the mammary glands of

wild-type, *Rag1*^{-/-}, *Cd4*^{-/-} and *Cd8*^{-/-} mice, all on the FVB/N background. Although the primary tumour volume was a little larger in *Cd4*^{-/-} mice and *Rag1*^{-/-} mice than in wild-type mice or *Cd8*^{-/-} mice (Supplementary Fig. 15a), pulmonary metastasis was greatly diminished in *Cd4*^{-/-} mice and *Rag1*^{-/-} mice but unaltered in *Cd8*^{-/-} mice compared with wild-type mice (Fig. 3a and Supplementary Fig. 15b). *ErbB2* mRNA quantification confirmed a threefold decrease in pulmonary metastasis in *Cd4*^{-/-} mice (Fig. 3b). MT2-cell-inoculated *Rag1*^{-/-} mice showed substantially reduced metastasis compared with MT2-cell-inoculated wild-type mice (Fig. 3c), but this was restored by transplanting CD4⁺ T cells, which were much more effective at restoring metastasis than CD8⁺ T cells or B cells (Fig. 3d). Treatment with exogenous RANKL restored pulmonary metastasis in the absence of T cells (Fig. 3e). CD4⁺CD25⁺ T cells were far more effective at restoring pulmonary metastasis in MT2-cell-inoculated *Rag1*^{-/-} mice than were CD4⁺CD25⁻ T cells (Fig. 3f), and this effect was almost completely blocked by administering RANK-Fc (Fig. 3g).

We examined whether RANKL stimulates the metastasis of human breast carcinoma cells. Of 19 carcinoma cell lines, the lines AU565, SK-BR-3, ZR-75-1 and SUM1315 expressed the most RANK protein or mRNA (Fig. 4a and Supplementary Fig. 16a). None of the cell lines expressed more *Rankl* mRNA than did untreated human mammary epithelial cells, which expressed RANKL only after treatment with progesterone (Supplementary Fig. 16b). ZR-75-1 and AU565 cells were inoculated into the number 2 mammary glands of nude mice (which lack mature T cells), and only ZR-75-1 cells formed slow growing primary tumours (data not shown), which gave rise to pulmonary metastases in 20% of mice, with an average multiplicity lower than 1 (Fig. 4b).

Administration of human RANKL increased metastasis incidence and multiplicity (Fig. 4b) and decreased maspin mRNA expression by ZR-75-1 cells (Fig. 4c). The carcinoma cell line T-47D, which is RANK⁻, did not respond to RANKL (Supplementary Fig. 17). In the stroma of human breast cancers, RANKL⁺ and FOXP3⁺ cells were present in close proximity to SMA⁺ cells (Fig. 4d). RANKL expression was higher in invasive ductal carcinomas than in ductal carcinomas *in situ* or lymph-node-positive tumours (Fig. 4e).

Our results provide a new mechanistic explanation for the association of CD4⁺ T-cell and T_{reg}-cell markers with a more aggressive behaviour in advanced breast cancers^{9,10,18,23}, by demonstrating that tumour-infiltrating CD4⁺CD25⁺FOXP3⁺ T_{reg} cells are a major source of RANKL, which stimulates the metastatic progression of RANK-expressing breast/mammary carcinoma cells (Supplementary Fig. 18). These findings and the demonstration that the pro-metastatic function of T cells can be replaced by exogenous RANKL differ from those obtained using the *MMTV-PyMT* mouse model of mammary carcinogenesis, in which CD4⁺ T helper 2 cells stimulate tumour progression through production of interleukin-4 and polarization of tumour-associated macrophages towards M2 macrophages¹⁸. Our mouse results correlate well with human breast cancer studies. In both cases, RANKL⁺ and FOXP3⁺ T_{reg} cells are concentrated in the tumour stroma and do not come into contact with carcinoma cells. Furthermore, in both cases, the presence of tumoral T_{reg} cells correlates with invasion, metastasis and poor prognosis^{9,10,23,24}. The recruitment of T_{reg} cells to CAFs is partly dependent on CCL5, a chemokine whose presence is associated with higher grades of human breast cancer and metastasis^{20,21}. Our suggestion that CAFs and CCL5 act through T cells,

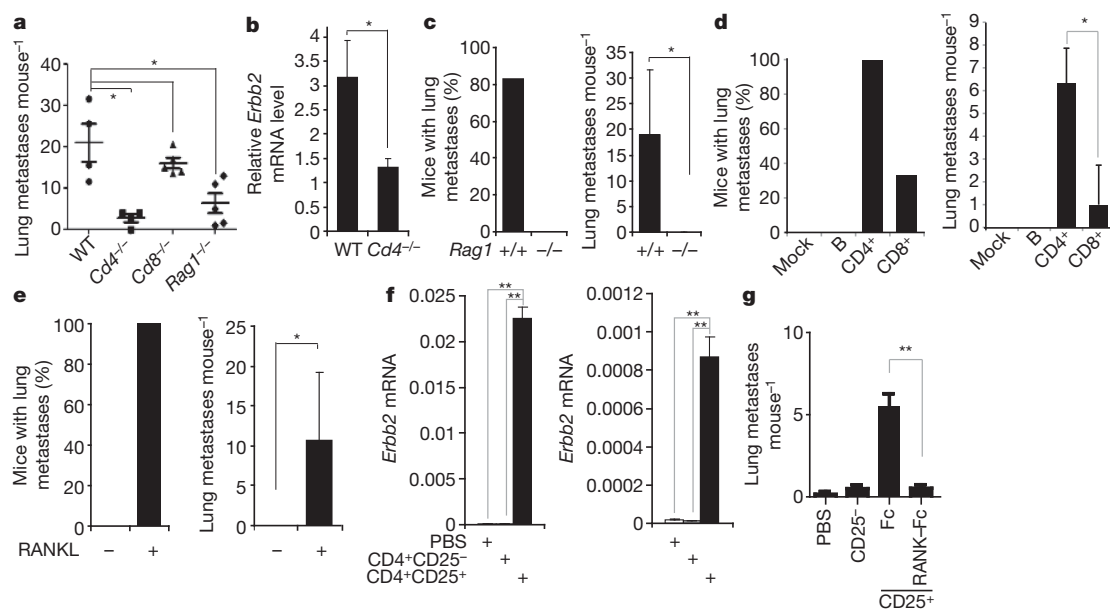


Figure 3 | Tumour-infiltrating CD4⁺ cells stimulate pulmonary metastasis. **a**, **b**, *MMTV-ErbB2*-transgenic PCaM cells (1×10^6 cells) were transplanted as in Fig. 1b to the indicated mouse strains. **a**, Lung metastasis multiplicity was calculated 56 days later. Individual mouse data are presented (mean, horizontal line; error bars, s.e.m.; $n = 4-6$). **b**, *ErbB2* expression in pulmonary metastases in the indicated mice was quantified by quantitative RT-PCR analyses of lung RNA with *MMTV-ErbB2*-specific primers. Data are presented as mean \pm s.e.m. ($n = 4$). **c**, MT2 cells were transplanted as in Fig. 1c to the indicated mouse strains. After 56 days, pulmonary metastasis incidence (left) and multiplicity (right) were quantified. Data are presented as mean \pm s.e.m. ($n = 6$). **d**, *Rag1*^{-/-} mice were reconstituted with the indicated cell types and then inoculated with MT2 cells 3 days after reconstitution. After 35 days, pulmonary metastasis incidence (left) and multiplicity (right) were quantified. Data are presented as mean \pm s.e.m. ($n = 3$). **e**, MT2-cell-inoculated *Rag1*^{-/-} mice were treated with BSA or RANKL ($80 \mu\text{g kg}^{-1}$) twice weekly, starting 1 week after

inoculation with tumour cells. After 35 days, pulmonary metastasis incidence (left) and multiplicity (right) were quantified. Data are presented as mean \pm s.e.m. ($n = 6-7$). **f**, MT2 cells were transplanted to *Rag1*^{-/-} mice, which were reconstituted 3 days before tumour cell injection with the indicated cell types or PBS as in **d**. After 56 days, *ErbB2* mRNA in lung metastases was measured by quantitative RT-PCR and normalized to cyclophilin mRNA levels. Each histogram represents an independent littermate group. Data are presented as mean \pm s.e.m. ($n = 3$). **g**, MT2 cells were transplanted to *Rag1*^{-/-} female littermates, which were reconstituted 3 days before tumour cell injection with PBS or T cells as in **f**. After 1 week, mice that had been reconstituted with CD4⁺CD25⁺ T cells received Fc protein (as a control) or RANK-Fc (2.5 mg kg^{-1}) twice weekly. Pulmonary metastasis multiplicity was measured on day 35 after inoculation with tumour cells. Data are presented as mean \pm s.e.m. ($n = 4-5$). *, $P < 0.05$; **, $P < 0.01$ (non-parametric Mann-Whitney *U* test).

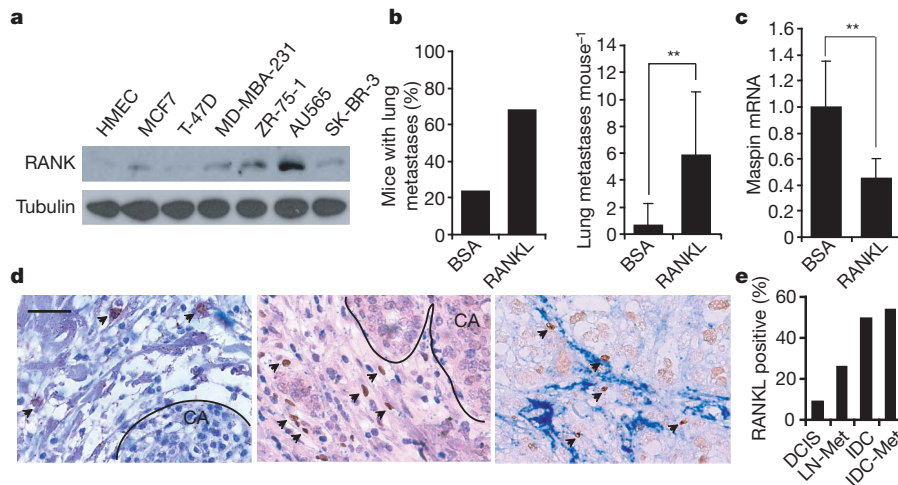


Figure 4 | RANKL in human breast cancer. **a**, RANK expression in the indicated human breast cancer cell lines was analysed by immunoblotting, with tubulin expression as a control. **b**, ZR-75-1 cells were transplanted to the number 2 mammary glands of nude mice, which were then treated with BSA or RANKL ($80 \mu\text{g kg}^{-1}$) twice weekly, starting 1 week after inoculation with tumour cells. After 84 days, pulmonary metastasis incidence (left) and multiplicity (right) were quantified. Data are presented as mean \pm s.e.m. ($n = 6$). **c**, Maspin mRNA in ZR-75-1 tumours from the mice in **b** was analysed by quantitative RT-PCR and normalized to cyclophilin mRNA levels. Data are presented as mean \pm s.e.m. ($n = 6$). **d**, Human breast cancer sections were

stained with anti-RANKL (left, brown), anti-FOXP3 (centre and right, brown), and anti-SMA (right, blue) antibodies. Arrows indicate RANKL⁺ cells (left) and FOXP3⁺ cells (centre and right). Scale bar, 20 μm (all images). **e**, Tissue arrays containing 50 samples each of ductal carcinoma *in situ* (DCIS), lymph-node macro-metastasis (LN-Met), invasive ductal carcinoma without metastasis (IDC) and IDC with metastasis (IDC-Met) were analysed for RANKL-expressing cells. Samples were considered to be positive when at least two cells per field ($\times 200$ magnification) were RANKL⁺. **b**, **c**, **, $P < 0.01$ (non-parametric Mann-Whitney U test).

in addition to having direct effects on carcinoma cells²⁰, fits the observation that the depletion of CAFs from mammary tumours decreased T_{reg}-cell infiltration and pulmonary metastasis²⁵. We postulate that the potent pro-metastatic effect of tumour-infiltrating T_{reg} cells can be dismantled by RANKL-RANK antagonists, which should leave the potential antitumorigenic activity of T helper 1 cells and CD8⁺ T cells intact. Such inhibitors of RANK signalling need to be used together with the surgical resection of primary breast tumours and the administration of antitumour immunotherapy to prevent the recurrence of metastatic disease, but may be an ineffective treatment after metastasis recurs. RANKL has also been reported to act at an early stage of tumour development, at which it mediates the tumour-promoting effect of progesterins^{26,27}. Thus, RANKL antagonists may also have a preventive ability.

METHODS SUMMARY

C57BL/6 *Rank*^{+/-} mice¹¹ were backcrossed to FVB/N *MMTV-ErbB2*-transgenic mice for at least six generations. FVB/N *MMTV-ErbB2/Rank*^{+/-} mice were intercrossed to generate FVB/N *MMTV-ErbB2/Rank*^{+/+} and FVB/N *MMTV-ErbB2/Rank*^{-/-} mice. *Cd4*^{-/-}, *Cd8*^{-/-} and *Rag1*^{-/-} mice were maintained on the FVB/N background. PCaM or MT2 single-cell suspensions were injected into the number 2 mammary gland. Tumour size was measured weekly with a caliper. At the study's end, mice were killed, and primary tumours and lungs were analysed.

Received 12 October; accepted 25 November 2010.

Published online 16 February; corrected 24 February 2011 (see full-text HTML version for details).

- Grivennikov, S. I., Greten, F. R. & Karin, M. Immunity, inflammation, and cancer. *Cell* **140**, 883–899 (2010).
- Luo, J. L. *et al.* Nuclear cytokine-activated IKK α controls prostate cancer metastasis by repressing Maspin. *Nature* **446**, 690–694 (2007).
- Fata, J. E. *et al.* The osteoclast differentiation factor osteoprotegerin-ligand is essential for mammary gland development. *Cell* **103**, 41–50 (2000).
- Cao, Y. *et al.* IKK α provides an essential link between RANK signaling and cyclin D1 expression during mammary gland development. *Cell* **107**, 763–775 (2001).
- Cao, Y., Luo, J. L. & Karin, M. I κ B kinase α kinase activity is required for self-renewal of ErbB2/Her2-transformed mammary tumor-initiating cells. *Proc. Natl Acad. Sci. USA* **104**, 15852–15857 (2007).
- Guy, C. T. *et al.* Expression of the *neu* protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. *Proc. Natl Acad. Sci. USA* **89**, 10578–10582 (1992).

- Slamon, D. J. *et al.* Studies of the HER-2/*neu* proto-oncogene in human breast and ovarian cancer. *Science* **244**, 707–712 (1989).
- Tiwari, R. K., Borgen, P. I., Wong, G. Y., Cordon-Cardo, C. & Osborne, M. P. HER-2/*neu* amplification and overexpression in primary human breast cancer is associated with early metastasis. *Anticancer Res.* **12**, 419–425 (1992).
- Bohling, S. D. & Allison, K. H. Immunosuppressive regulatory T cells are associated with aggressive breast cancer phenotypes: a potential therapeutic target. *Mod. Pathol.* **21**, 1527–1532 (2008).
- Ohara, M. *et al.* Possible involvement of regulatory T cells in tumor onset and progression in primary breast cancer. *Cancer Immunol. Immunother.* **58**, 441–447 (2009).
- Dougall, W. C. *et al.* RANK is essential for osteoclast and lymph node development. *Genes Dev.* **13**, 2412–2424 (1999).
- Body, J. J. *et al.* Effects of denosumab in patients with bone metastases, with and without previous bisphosphonate exposure. *J. Bone Miner. Res.* **25**, 440–446 (2010).
- Fizazi, K. *et al.* Randomized phase II trial of denosumab in patients with bone metastases from prostate cancer, breast cancer, or other neoplasms after intravenous bisphosphonates. *J. Clin. Oncol.* **27**, 1564–1571 (2009).
- Ercolini, A. M. *et al.* Recruitment of latent pools of high-avidity CD8⁺ T cells to the antitumor immune response. *J. Exp. Med.* **201**, 1591–1602 (2005).
- Hsu, H. *et al.* Tumor necrosis factor receptor family member RANKL mediates osteoclast differentiation and activation induced by osteoprotegerin ligand. *Proc. Natl Acad. Sci. USA* **96**, 3540–3545 (1999).
- Jones, D. H. *et al.* Regulation of cancer cell migration and bone metastasis by RANKL. *Nature* **440**, 692–696 (2006).
- Kawai, T. *et al.* B and T lymphocytes are the primary sources of RANKL in the bone resorptive lesion of periodontal disease. *Am. J. Pathol.* **169**, 987–998 (2006).
- DeNardo, D. G. *et al.* CD4⁺ T cells regulate pulmonary metastasis of mammary carcinomas by enhancing protumor properties of macrophages. *Cancer Cell* **16**, 91–102 (2009).
- Fontenot, J. D., Gavin, M. A. & Rudensky, A. Y. Foxp3 programs the development and function of CD4⁺CD25⁺ regulatory T cells. *Nature Immunol.* **4**, 330–336 (2003).
- Karnoub, A. E. *et al.* Mesenchymal stem cells within tumour stroma promote breast cancer metastasis. *Nature* **449**, 557–563 (2007).
- Luboshits, G. *et al.* Elevated expression of the CC chemokine regulated on activation, normal T cell expressed and secreted (RANTES) in advanced breast carcinoma. *Cancer Res.* **59**, 4681–4687 (1999).
- Gao, J. L. *et al.* Structure and functional expression of the human macrophage inflammatory protein 1 α /RANTES receptor. *J. Exp. Med.* **177**, 1421–1427 (1993).
- Gobert, M. *et al.* Regulatory T cells recruited through CCL22/CCR4 are selectively activated in lymphoid infiltrates surrounding primary breast tumors and lead to an adverse clinical outcome. *Cancer Res.* **69**, 2000–2009 (2009).
- Bates, G. J. *et al.* Quantification of regulatory T cells enables the identification of high-risk breast cancer patients and those at risk of late relapse. *J. Clin. Oncol.* **24**, 5373–5380 (2006).

25. Liao, D., Luo, Y., Markowitz, D., Xiang, R. & Reisfeld, R. A. Cancer associated fibroblasts promote tumor growth and metastasis by modulating the tumor immune microenvironment in a 4T1 murine breast cancer model. *PLoS ONE* **4**, e7965 (2009).
26. Gonzalez-Suarez, E. *et al.* RANK ligand mediates progestin-induced mammary epithelial proliferation and carcinogenesis. *Nature* **468**, 103–107 (2010).
27. Schramek, D. *et al.* Osteoclast differentiation factor RANKL controls development of progestin-driven mammary cancer. *Nature* **468**, 98–102 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Coussens for discussions and suggestions, V. Temkin and G. He for advice, and H. Herschman for critical reading. Work was supported by the National Institutes of Health (NIH). W.T. and W.Z. were supported by postdoctoral fellowships from the Susan G. Komen for the Cure. A.S. was supported by NIH Asthma

Research and Cancer Therapeutic training grants, and S.G. was supported by a research fellowship from the Crohn's and Colitis Foundation of America. M.K. is an American Cancer Society research professor.

Author Contributions W.T., W.Z., A.S. and M.K. designed the research. W.T., W. Z. and A.S. performed most of the experiments and analysed the data. W.T., W.Z., A.S. and M. K. wrote the manuscript. S.G. performed some of the experiments and analysed their data. J.Q.C. and R.M.H. provided essential reagents, *in vivo* imaging and technical discussions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.K. (karinoffice@ucsd.edu).

Co-directional replication–transcription conflicts lead to replication restart

Houra Merrikh^{1*}, Cristina Machón^{2*}, William H. Grainger², Alan D. Grossman¹ & Panos Soultanas²

Head-on encounters between the replication and transcription machineries on the lagging DNA strand can lead to replication fork arrest and genomic instability^{1,2}. To avoid head-on encounters, most genes, especially essential and highly transcribed genes, are encoded on the leading strand such that transcription and replication are co-directional. Virtually all bacteria have the highly expressed ribosomal RNA genes co-directional with replication³. In bacteria, co-directional encounters seem inevitable because the rate of replication is about 10–20-fold greater than the rate of transcription. However, these encounters are generally thought to be benign^{2,4–9}. Biochemical analyses indicate that head-on encounters¹⁰ are more deleterious than co-directional encounters⁸ and that in both situations, replication resumes without the need for any auxiliary restart proteins, at least *in vitro*. Here we show that *in vivo*, co-directional transcription can disrupt replication, leading to the involvement of replication restart proteins. We found that highly transcribed rRNA genes are hotspots for co-directional conflicts between replication and transcription in rapidly growing *Bacillus subtilis* cells. We observed a transcription-dependent increase in association of the replicative helicase and replication restart proteins where head-on and co-directional conflicts occur. Our results indicate that there are co-directional conflicts between replication and transcription *in vivo*. Furthermore, in contrast to the findings *in vitro*, the replication restart machinery is involved *in vivo* in resolving potentially deleterious encounters due to head-on and co-directional conflicts. These conflicts probably occur in many organisms and at many chromosomal locations and help to explain the presence of important auxiliary proteins involved in replication restart and in helping to clear a path along the DNA for the replisome.

The DNA replication machinery (replisome) often encounters obstacles along the genome that can cause replication fork arrest^{1,2} (Supplementary Fig. 1). In bacteria, replication, transcription and translation occur concurrently, and RNA polymerase (RNAP) transcribing the lagging strand (head-on relative to replication) is a well-known obstacle encountered by the replisome^{1,4–7,9,11–14}. Transcription–replication conflicts are compounded during rapid growth when transcription initiation of many genes, especially those encoding the protein synthesis machinery, increases. In *Bacillus subtilis*, head-on conflicts between replication and transcription slow the overall rate of replication fork progression, largely owing to obstruction of the replisome^{9,12}.

rRNA genes (Supplementary Fig. 2) are among the most highly expressed in rapidly growing bacteria, and are co-directional with replication (that is, encoded on the leading strand), thereby avoiding head-on conflicts³. Nonetheless, co-directional encounters between bacterial replication and transcription machineries seem inevitable because the rate of replication ($\sim 500\text{--}1,000$ nucleotides s^{-1}) is $\sim 10\text{--}20$ times faster than that of transcription¹. The potential for co-directional conflicts is widely recognized^{1,2,11}, but these conflicts have not been detected *in vivo*^{4–6,9}, with the exception of co-directionally positioned transcription

terminators that can inhibit replication fork progression⁷. In addition, during co-directional encounters engineered to occur *in vitro*, the replicative helicase translocating along the lagging strand simply displaces RNAP translocating along the leading strand^{2,8}. Thus, co-directional encounters are generally thought to have little or no effect on replication^{1,2,11}.

All organisms have mechanisms for loading a helicase onto DNA during replication fork assembly. DnaA-dependent mechanisms load the replicative helicase at the origin of replication, *oriC*, and recombination-based and PriA-dependent mechanisms restart forks from stalled sites¹⁵. Although DnaA and PriA are ubiquitous, other helicase-loading proteins differ among bacteria. In *B. subtilis*, and other low G+C Gram-positive organisms, DnaD and DnaB participate in loading the replicative helicase, DnaC, both at *oriC* and during replication restart at stalled forks^{16–20}. We measured association of DnaD, DnaB and helicase with chromosomal regions using chromatin immunoprecipitation (ChIP) and either quantitative real time polymerase chain reaction (ChIP–qPCR) to detect specific regions, or hybridization to DNA microarrays (ChIP–chip) for genome-wide analysis (Methods).

We analysed head-on conflicts between transcription and replication in a specific chromosomal region. *Pxis*, a promoter from the conjugative transposon ICEBsI, is highly expressed in the absence of the transposon-encoded immunity repressor ImmR (Methods). Using ChIP–qPCR, we found that there was a twofold increase in association of DnaD, DnaB and helicase with the chromosomal region (*thrC*) expressing a *Pxis-lacZ* fusion compared with other chromosomal regions (Supplementary Fig. 3). In contrast, when *Pxis-lacZ* was off, in cells containing ICEBsI and its repressor, there was no detectable enrichment of these proteins (Supplementary Fig. 3). Thus, head-on conflicts between replication and transcription *in vivo* can be detected by increased association of the replicative helicase and the replication restart proteins with the region of conflict. The association of helicase probably indicates replisome stalling in this region. Association of DnaD and DnaB indicates that these proteins are probably acting to reload the helicase for replication restart. It is formally possible that DnaD and DnaB are part of the replisome or are associated with the replication fork. If true, then their association could indicate fork stalling and/or restart. However, neither DnaD nor DnaB are required for replication elongation, nor do they seem to be associated with the replication fork^{21,22}. Thus, it seems most likely that their association is indicative of replication restart.

We also detected the head-on conflict between replication and transcription in ChIP–chip assays. When *Pxis-lacZ* was expressed, there was increased association of DnaB with this region (Fig. 1a). In contrast, in cells not expressing *Pxis-lacZ*, there was little or no detectable association of DnaB with this region (Fig. 1b), although there was association of DnaB with other chromosomal regions (see later). These results indicate that association of DnaB with the region near *Pxis-lacZ* depends on transcription from *Pxis*.

We analysed genome-wide association of the replication restart proteins DnaD and DnaB in wild-type cells using ChIP–chip. There

¹Department of Biology, Building 68-530, M.I.T., Cambridge, Massachusetts 02139, USA. ²Centre for Biomolecular Sciences, School of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK.

*These authors contributed equally to this work.

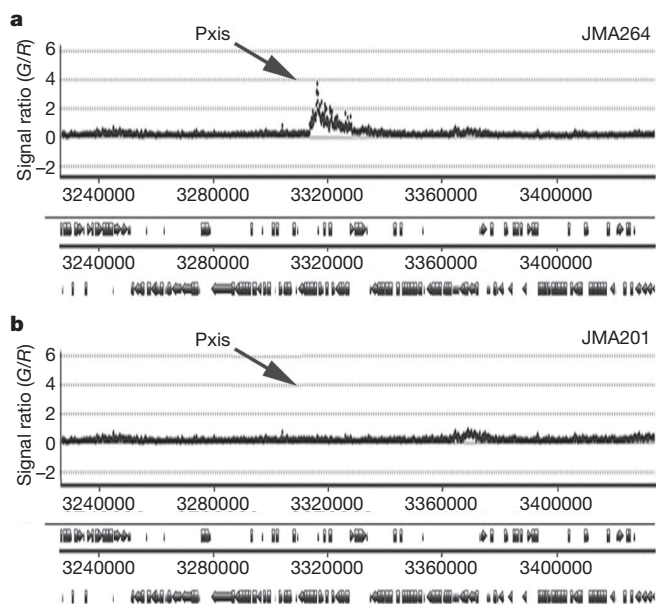


Figure 1 | Head-on conflicts between transcription and replication cause increased association of helicase loader protein DnaB. **a, b,** Association of DnaB was assessed in ChIP-chip experiments in strains containing Pxis-*lacZ* inserted at *thrC*. Cells were grown in rich medium (Luria broth (LB)) and sampled during exponential growth. The relative enrichment, indicated as the signal ratio (G/R) of the sample (green channel) divided by the control (red channel), of a given chromosomal position is plotted on the y-axis versus the chromosomal position on the x-axis (in bp clockwise from *oriC*). Data are shown for the chromosomal region from ~3,240 kb to ~3,400 kb. The location of Pxis-*lacZ* inserted at *thrC* is indicated. Pxis-*lacZ* is head-on with replication. Coding sequences are indicated with arrows below the graphs. **a,** Data from cells expressing Pxis-*lacZ* (strain JMA264). **b,** Data from cells not expressing Pxis-*lacZ* (strain JMA201). These findings were verified by qPCR (Supplementary Fig. 3).

was significant enrichment of the *oriC* region and the ten *rrn* (rRNA) loci in the DnaD (Supplementary Fig. 4a) and DnaB (Fig. 2a and Supplementary Fig. 5) immunoprecipitates compared to most other chromosomal regions (Supplementary Discussion). *rrn* loci are among the most highly transcribed genes during rapid growth and are transcribed on the leading strand. The presence of DnaD and DnaB might

be indicative of replication restart after fork stalling in these highly transcribed regions. This enrichment was dependent on rapid growth. During slow growth in minimal medium, there was little or no detectable enrichment of *rrn* loci in the DnaD (Supplementary Fig. 4b) or DnaB (Fig. 2b and Supplementary Fig. 5) immunoprecipitates. Because the genome sequencing project for *B. subtilis* used a 'consensus sequence' for all rRNA genes²³, and the sequences of each were reported as identical, our results indicate that at least one, and probably several *rrn* loci are enriched in the immunoprecipitates (see later), and that this enrichment is reproducible and most noticeable during rapid growth when the *rrn* loci are most highly expressed.

We also used ChIP-qPCR to measure association of DnaD and DnaB with *rrn* loci. Primer pairs were designed to detect DNA from three different regions of rRNA genes (Supplementary Fig. 2b–d): (1) a region unique to *rrnD* immediately upstream of its 16S gene and far from *oriC* (Supplementary Fig. 2a, b); (2) a region just upstream and overlapping the 16S genes of *rrnO*, *E*, *D* and *B* (Supplementary Fig. 2c); and (3) a region that should be common to all 23S genes (Supplementary Fig. 2d). There was significant enrichment of the *rrn* loci in samples from both the DnaD (Fig. 3a) and DnaB (Fig. 3b) immunoprecipitates, similar to the results from the ChIP-chip analyses (Fig. 2 and Supplementary Fig. 4). The ChIP signals were similar with all three probes because the qPCR normalizes to the number of copies of each region. These results indicate that DnaD and DnaB are associated with *rrnD*, and probably most or all *rrn* loci.

Even though co-directional conflicts between replication and transcription are not thought to be deleterious to replication^{2,4–7}, the association of DnaD and DnaB with *rrn* loci is probably due to replication fork stalling and restart. If true, then this association should depend on transcription, and the replicative helicase should also be associated with *rrn* loci. There was a decrease in association of DnaD (Fig. 3a) and DnaB (Fig. 3b) with *rrn* loci after inhibition of transcription initiation with rifampicin (which blocks RNAP at the promoter). In addition, the replicative helicase was associated with *rrn* loci during rapid growth (Fig. 4), and this association also decreased following treatment with rifampicin (Fig. 4). Using strains in which replication initiates from an ectopic origin inserted near *oriC* to maintain proper orientation of transcription and replication (Supplementary Discussion), we found that association of helicase at the *rrn* loci was independent of DnaA and replication initiation from *oriC* (Fig. 4). Together, these results support the hypothesis that association of helicase, DnaD and DnaB with the *rrn* loci is a consequence of replication fork stalling and

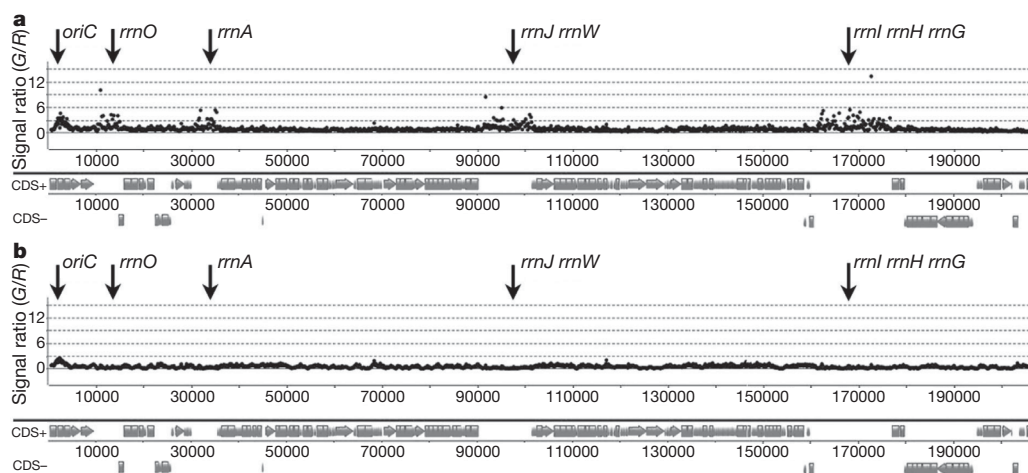


Figure 2 | ChIP-chip analysis of DnaB. **a, b,** Wild-type cells (strain 168) were grown in LB (a) or defined minimal medium (b) and sampled during exponential growth. The coding sequences (CDS) are indicated with arrows below the graph. Data are plotted as in Fig. 1, except the chromosomal positions are shown from 0 kb (*oriC*) to just past *rrnI*, *H* and *G* at ~200 kb. Similar results were obtained at each identical *rrn* with both DnaD and DnaB, indicating the

reproducibility of the data. Results were also confirmed by qPCR with independent samples from different strains (Fig. 3). Data from other *rrn* regions are presented in Supplementary Fig. 5. The *rrn* sequences represent a consensus and were thus presented as identical²³. For clarity and simplicity, we unambiguously label each individual locus according to its chromosomal location.

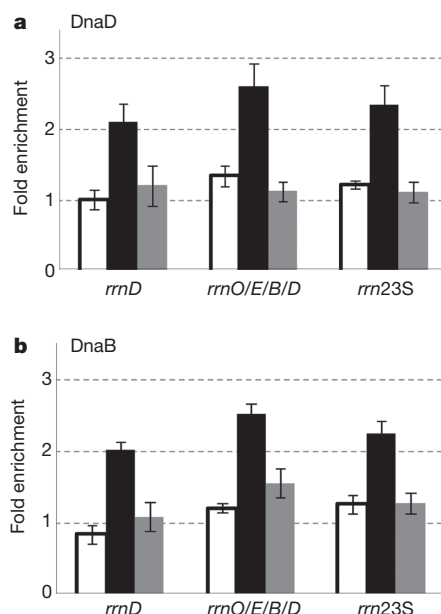


Figure 3 | Association of helicase loader proteins DnaD and DnaB with *rrn* loci depends on transcription and the replication restart protein *priA*.

a, b, Wild-type cells (AG174) and the *priA-ssrA** mutant (WKS338) were grown to mid-exponential phase in LB medium. For wild type, samples were taken in the absence of (black bars) or 4 min after treatment (grey bars) with rifampicin ($30 \mu\text{g ml}^{-1}$) to block transcription initiation. The *priA-ssrA** mutant (grown in the presence of $1 \mu\text{g ml}^{-1}$ of chloramphenicol to maintain selection for the mutant allele) was sampled in the absence of rifampicin (white bars). Association of DnaD (**a**) and DnaB (**b**) was analysed by ChIP-qPCR with three different primer pairs (Supplementary Fig. 2) that recognize the indicated *rrn* loci. The ChIP-qPCR signals are normalized to gene copy number (Methods), so that the signal for the 23S *rrn* probe, which should detect all 10 *rrn* loci, is normalized per locus. Data are averages from at least three independent cultures. Error bars represent standard error.

restart due to co-directional conflicts between replication and transcription. These results, and the finding that association of DnaD, DnaB and helicase with the *rrn* loci was dependent on rapid growth, indicate that a high density of elongating RNAP molecules cause replication fork stalling and restart. We estimate that there are at least 40, and probably >100, RNAP molecules per rRNA operon in *B. subtilis* during rapid growth (Supplementary Discussion).

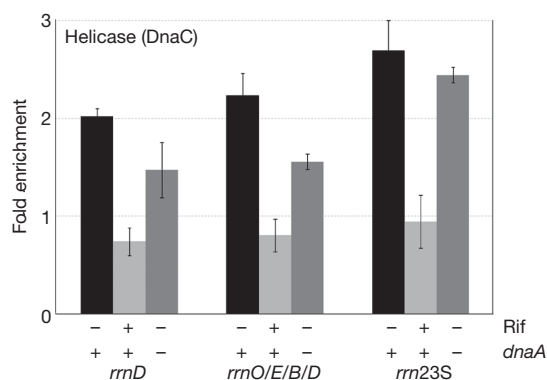


Figure 4 | Association of the replicative helicase with *rrn* loci depends on transcription and is independent of *dnaA*. Samples from wild-type cells (AG174) with and without rifampicin (Rif) and a *dnaA* null mutant (AIG200) were grown and analysed as described for Fig. 3. Data are averages from at least three independent cultures. Error bars represent standard error. We also tested association of DnaB with the *rrn* loci using the 23S *rrn* probe and found similar association in the *dnaA* null mutant (data not shown).

The essential replication restart protein PriA was required for association of DnaD and DnaB with *rrn* loci (Fig. 3). PriA enables the resumption of replication at regions of replication fork collapse^{15–17,19,24–26}. *priA* mutants interact genetically with mutations that affect RNAP progression and stability^{27,28}, indicating a possible role for the restart machinery in resolving conflicts between transcription and replication *in vivo*. In a partially defective *priA* mutant (Methods), association of DnaD and DnaB with the *rrn* regions was reduced (Fig. 3). These results strongly support the conclusion that PriA, DnaD and DnaB are functioning in replication restart at rRNA genes (Supplementary Fig. 1).

In vitro studies with purified *E. coli* proteins indicate that during both co-directional and head-on encounters between the replisome and RNAP, RNAP is displaced and the replisome resumes replication without coming off the DNA, without the need for replication restart proteins^{8,10}. It is not clear if or how frequently this happens *in vivo* in rapidly growing cells where the replisome probably encounters multiple RNAP molecules aligned in tandem at highly transcribed genes⁴. Cells have mechanisms for removing RNAP to allow progression of replication forks^{13,14,27,29}, thereby avoiding such conflicts.

Our findings indicate that *in vivo*, both head-on and co-directional encounters between replication and transcription can lead to replication fork stalling and recruitment of the helicase loading machinery. Head-on encounters are clearly more severe as inversion of *rrn* operons causes an appreciable slowing of replication^{4–6,9}. Helicase loading machineries are probably used in all organisms to restart replication in regions of both head-on and co-directional transcription–replication conflicts. During these conflicts, the helicase will sometimes disengage from the template DNA, leaving behind a forked DNA substrate with a single-stranded region on the lagging strand. PriA binds strongly to this type of substrate. This is probably followed by the sequential recruitment of *B. subtilis* DnaD and DnaB, and then DnaI-mediated loading of the replicative helicase¹⁷ (Supplementary Fig. 1).

We estimate that ~5–10% of cells in an asynchronous population have a conflict between the transcription and replication machineries at an rRNA operon. This estimate is based on a ~50–100-fold greater association of helicase at *oriC* in a synchronous population at the time of replication initiation²⁰ than at one of the ten rRNA operons, assuming that there are similar crosslinking efficiencies at *oriC* and *rrn* loci. The co-directional conflicts between replication and transcription probably account for a significant fraction of endogenous events requiring repair of stalled replication forks^{1,26}, and may even account for some of the sensitivity to rapid growth conditions of *priA* mutants defective in replication restart^{25,30}. Because inability to repair a stalled fork would prevent completion of a replication cycle and the production of viable progeny, there is strong selective pressure to avoid such catastrophes.

METHODS SUMMARY

Strains are listed in the Supplementary Table. Relevant properties are described in the text. Strain constructions, growth conditions and oligonucleotides are described in Methods. Standard procedures were used for ChIP experiments and are described in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 February; accepted 9 December 2010.

1. Mirkin, E. V. & Mirkin, S. M. Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.* **71**, 13–35 (2007).
2. Pomerantz, R. T. & O'Donnell, M. What happens when replication and transcription complexes collide? *Cell Cycle* **9**, 2535–2541 (2010).
3. Rocha, E. P. The replication-related organization of bacterial genomes. *Microbiology* **150**, 1609–1627 (2004).
4. French, S. Consequences of replication fork movement through transcription units *in vivo*. *Science* **258**, 1362–1365 (1992).
5. Olavarrieta, L., Hernandez, P., Krimer, D. B. & Schwartzman, J. B. DNA knotting caused by head-on collision of transcription and replication. *J. Mol. Biol.* **322**, 1–6 (2002).
6. Mirkin, E. V. & Mirkin, S. M. Mechanisms of transcription–replication collisions in bacteria. *Mol. Cell. Biol.* **25**, 888–895 (2005).

7. Mirkin, E. V., Castro Roa, D., Nudler, E. & Mirkin, S. M. Transcription regulatory elements are punctuation marks for DNA replication. *Proc. Natl Acad. Sci. USA* **103**, 7276–7281 (2006).
8. Pomerantz, R. T. & O'Donnell, M. The replisome uses mRNA as a primer after colliding with RNA polymerase. *Nature* **456**, 762–766 (2008).
9. Srivatsan, A., Tehranchi, A., MacAlpine, D. M. & Wang, J. D. Co-orientation of replication and transcription preserves genome integrity. *PLoS Genet.* **6**, e1000810 (2010).
10. Pomerantz, R. T. & O'Donnell, M. Direct restart of a replication fork stalled by a head-on RNA polymerase. *Science* **327**, 590–592 (2010).
11. Rudolph, C. J., Dhillon, P., Moore, T. & Lloyd, R. G. Avoiding and resolving conflicts between DNA replication and transcription. *DNA Repair* **6**, 981–993 (2007).
12. Wang, J. D., Berkmen, M. B. & Grossman, A. D. Genome-wide coorientation of replication and transcription reduces adverse effects on replication in *Bacillus subtilis*. *Proc. Natl Acad. Sci. USA* **104**, 5608–5613 (2007).
13. Tehranchi, A. K. *et al.* The transcription factor DksA prevents conflicts between DNA replication and transcription machinery. *Cell* **141**, 595–605 (2010).
14. Boubakri, H., de Septenville, A. L., Viguera, E. & Michel, B. The helicases DinG, Rep and UvrD cooperate to promote replication across transcription units *in vivo*. *EMBO J.* **29**, 145–157 (2010).
15. Heller, R. C. & Marians, K. J. Replisome assembly and the direct restart of stalled replication forks. *Nature Rev. Mol. Cell Biol.* **7**, 932–943 (2006).
16. Bruand, C., Farache, M., McGovern, S., Ehrlich, S. D. & Polard, P. DnaB, DnaD and DnaI proteins are components of the *Bacillus subtilis* replication restart primosome. *Mol. Microbiol.* **42**, 245–256 (2001).
17. Marsin, S., McGovern, S., Ehrlich, S. D., Bruand, C. & Polard, P. Early steps of *Bacillus subtilis* primosome assembly. *J. Biol. Chem.* **276**, 45818–45825 (2001).
18. Rokop, M. E., Auchtung, J. M. & Grossman, A. D. Control of DNA replication initiation by recruitment of an essential initiation protein to the membrane of *Bacillus subtilis*. *Mol. Microbiol.* **52**, 1757–1767 (2004).
19. Bruand, C. *et al.* Functional interplay between the *Bacillus subtilis* DnaD and DnaB proteins essential for initiation and re-initiation of DNA replication. *Mol. Microbiol.* **55**, 1138–1150 (2005).
20. Smits, W. K., Goranov, A. I. & Grossman, A. D. Ordered association of helicase loader proteins with the *Bacillus subtilis* origin of replication *in vivo*. *Mol. Microbiol.* **75**, 452–461 (2010).
21. Imai, Y. *et al.* Subcellular localization of Dna-initiation proteins of *Bacillus subtilis*: evidence that chromosome replication begins at either edge of the nucleoids. *Mol. Microbiol.* **36**, 1037–1048 (2000).
22. Meile, J. C., Wu, L. J., Ehrlich, S. D., Errington, J. & Noiro, P. Systematic localisation of proteins fused to the green fluorescent protein in *Bacillus subtilis*: identification of new proteins at the DNA replication factory. *Proteomics* **6**, 2135–2146 (2006).
23. Kunst, F. *et al.* The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249–256 (1997).
24. McGlynn, P., Al-Deib, A. A., Liu, J., Marians, K. J. & Lloyd, R. G. The DNA replication protein PriA and the recombination protein RecG bind D-loops. *J. Mol. Biol.* **270**, 212–221 (1997).
25. Polard, P. *et al.* Restart of DNA replication in Gram-positive bacteria: functional characterisation of the *Bacillus subtilis* PriA initiator. *Nucleic Acids Res.* **30**, 1593–1605 (2002).
26. Gabbai, C. B. & Marians, K. J. Recruitment to stalled replication forks of the PriA DNA helicase and replisome-loading activities is essential for survival. *DNA Repair* **9**, 202–209 (2010).
27. Trautinger, B. W., Jaktaji, R. P., Rusakova, E. & Lloyd, R. G. RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. *Mol. Cell* **19**, 247–258 (2005).
28. Mahdi, A. A., Buckman, C., Harris, L. & Lloyd, R. G. Rep and PriA helicase activities prevent RecA from provoking unnecessary recombination during replication fork repair. *Genes Dev.* **20**, 2135–2147 (2006).
29. Guy, C. P. *et al.* Rep provides a second motor at the replisome to promote duplication of protein-bound DNA. *Mol. Cell* **36**, 654–666 (2009).
30. Nurse, P., Zavitz, K. H. & Marians, K. J. Inactivation of the *Escherichia coli* priA DNA replication protein induces the SOS response. *J. Bacteriol.* **173**, 6686–6693 (1991).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Grainger, C. Lee, T. Baker, and W. K. Smits for discussions, W. K. Smits for constructing the *priA-ssrA** mutant, and C. Lee, C. Bonilla, S. P. Bell, and J. D. Wang for comments on the manuscript. Work in the P.S. laboratory was supported by Biotechnology and Biological Sciences Research Council grant BB/E006450/1 and a Wellcome Trust grant 091968/Z/10/Z. Work in the A.D.G. laboratory was supported by NIH grant GM41934 and H.M. was supported in part by NIH postdoctoral fellowship GM093408. The Biotechnology and Biological Sciences Research Council and the Royal Society provided funds for a sabbatical visit of P.S. in the A.D.G. laboratory.

Author Contributions H.M., C.M., W.H.G., A.D.G. and P.S. designed the research and analysed the results; H.M., C.M. and W.H.G. performed the experiments; H.M., A.D.G. and P.S. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.D.G. (adg@mit.edu) or P.S. (Panos.Soultanas@nottingham.ac.uk).

METHODS

Strains. *B. subtilis* strain 168 (*trp*) and derivatives of strain JH642 (*trp phe*) were used for all experiments (Supplementary Table) and were constructed by standard procedures³¹. The *priA* mutation was constructed by attaching an *ssrA** tag onto the 3' end of *priA*. *ssrA** encodes a tag that makes the gene product unstable in the presence of the adaptor protein SspB³². A PCR product carrying a C-terminal fragment of *priA* was cloned into pKG1268 to give the plasmid pGCS-*priA*. This plasmid was introduced by single crossover into *priA*, generating *priA-ssrA**, in strain KG1098 (*amyE::Pspank(-7TA)-sspB, spc*) that contains *sspB* under control of the weakened IPTG-inducible promoter *Pspank(-7TA)*³². The *priA-ssrA** mutant (WKS338) was defective even in the absence of induction of SspB expression, probably because of the low level of expression without induction.

Media and growth conditions. For all experiments, cells were grown at 30 °C and samples taken during mid-exponential phase. Growth was in either rich medium (LB) or LeMaster minimal medium³³, prepared by suspending: L-Ala 0.5 g, L-Arg(HCl) 0.58 g, L-Asp 0.41 g, L-Cys 0.03 g, L-Glu 0.67 g, L-Gly 0.54 g, L-His 0.06 g, L-Ile 0.23 g, L-Leu 0.23 g, L-Lys(HCl) 0.42 g, L-Met 0.5 g, L-Phe 0.13 g, L-Pro 0.10 g, L-Ser 2.08 g, L-Thr 0.23 g, L-Tyr 0.17 g, L-Val 0.23 g, adenine 0.5 g, guanosine 0.67 g, thymine 0.17 g, uracil 0.5 g, sodium acetate 1.50 g, succinic acid 1.50 g, ammonium chloride 0.75 g, sodium hydroxide 1.08 g, and anhydrous K₂HPO₄·3H₂O 8 g in one litre of distilled H₂O and autoclaving. The pH of this pre-medium was checked and adjusted to ~7.5 if necessary. The final medium was completed by the addition of filtered-sterilized glucose (10 g 100 ml⁻¹), MgSO₄·7H₂O (0.25 g 100 ml⁻¹), FeSO₄ (4.2 mg 100 ml⁻¹), thiamine-HCl (5 mg 100 ml⁻¹) and concentrated HCl (8 µl 100 ml⁻¹).

ChIP-chip analysis. Polyclonal rabbit anti-DnaB, anti-DnaD and anti-DnaI antibodies were produced and tested as described³⁴. Preparation of DNA samples for ChIP-chip analysis was carried out as described³⁴ with minor modifications. An overnight culture of *B. subtilis* (strain 168) was used to inoculate 800 ml of LB or LeMaster minimal medium. The culture was incubated at 30 °C and during exponential growth (OD_{595 nm} = 0.8) 1% v/v formaldehyde was added for 20 min to crosslink protein–DNA complexes. The reaction was quenched by 0.5 M glycine. Preparation of samples for microarray analysis was carried out as described³⁴.

B. subtilis (strain 168) Agilent 4x44K ChIP arrays with AMADID 023001 were prepared by Oxford Gene Technologies (OGT) who also carried out array hybridizations and provided the final data. Each array comprised 41,770 probes in total, covering 4,185 genes. Each probe was 60 bp and generated using Agilent's inkjet *in situ* synthesis technology. The probes covered comprehensively the entire genome. They had an average spacing of ~100 bp with a maximum interprobe distance of

~140 bp. The reference sample in the red (Cy5) channel was genomic *B. subtilis* (strain 168) DNA. Data analysis was carried out with a ChIP browser developed and supplied by OGT.

ChIP and quantitative real time PCRs. Cells were grown in LB medium at 30 °C to mid-exponential phase. Samples were crosslinked as above and rabbit polyclonal antibodies against DnaD, DnaB and helicase (DnaC) were used as described previously²⁰. Immunoprecipitations were done at room temperature (approximately 22 °C) for 2 h with the antibody, followed by 1 h with 3% protein A-sepharose beads.

The quantitative real-time PCRs were performed as described²⁰. Primer pairs included: HM84 (5'-CAAGCTCACAGCGGCGGAAAAT-3') and HM85 (5'-GCCCTAGTTTGACTGACTACGC-3') that amplify a sequence upstream of *rrnD*; HM43 (5'-CTGCACGACGACAGGTACACACAGTG-3') and HM44 (5'-CTCCCATCTGTCCGCTCGACTTGC-3') that amplify sequences beginning upstream of *rrnO*, *rrnE*, *rrnD* and *rrnB* and extending into the 16S rRNA gene; HM80 (5'-AGGATAGGGTAAGCGCGGTATT-3') and HM81 (5'-TTCTCTCGATCACCTTAGGATTTC-3') that amplify sequences internal to all 23S rRNA genes.

yhaX is a chromosomal locus that does not have increased association with DnaD, DnaB and helicase and was used for comparison. *yhaX* was detected with primers WKS145 (5'-CGAGCAAGGTGTCGCTTA-3') and WKS146 (5'-GCAGCGGTCATCATGTA-3').

RT-qPCRs were quantified by comparison of the crossing-point values generated in the PCR for each sample to standard curves generated for that primer set using chromosomal DNA as template. Data were first normalized to immunoprecipitations of *yhaX*, and then to gene copy number as determined by PCRs from 'total' samples (lysates pre-immunoprecipitation). The final fold enrichment was determined as: (*x* IP/*yhaX* IP) / (*x* total/*yhaX* total), where *x* represents the region of interest. All data presented are the averages of at least 3 biological replicates ± standard error.

- Harwood, C. R. & Cutting, S. M. *Molecular Biological Methods for Bacillus* (John Wiley & Sons, 1990).
- Griffith, K. L. & Grossman, A. D. Inducible protein degradation in *Bacillus subtilis* using heterologous peptide tags and adaptor proteins to target substrates to the protease ClpXP. *Mol. Microbiol.* **70**, 1012–1025 (2008).
- LeMaster, D. M. & Richards, F. M. 1H-15N heteronuclear NMR studies of *Escherichia coli* thioredoxin in samples isotopically labeled by residue type. *Biochemistry* **24**, 7263–7268 (1985).
- Grainger, W. H., Machon, C., Scott, D. J. & Soultanas, P. DnaB proteolysis *in vivo* regulates oligomerization and its localization at *oriC* in *Bacillus subtilis*. *Nucleic Acids Res.* **38**, 2851–2864 (2010).

Crystal structure of the CusBA heavy-metal efflux complex of *Escherichia coli*

Chih-Chia Su^{1*}, Feng Long^{1*}, Michael T. Zimmermann², Kanagalaghatta R. Rajashankar³, Robert L. Jernigan^{2,4} & Edward W. Yu^{1,2,4,5}

Gram-negative bacteria, such as *Escherichia coli*, expel toxic chemicals through tripartite efflux pumps that span both the inner and outer membrane. The three parts are an inner membrane, substrate-binding transporter; a membrane fusion protein; and an outer-membrane-anchored channel. The fusion protein connects the transporter to the channel within the periplasmic space. A crystallographic model of this tripartite efflux complex has been unavailable because co-crystallization of the various components of the system has proven to be extremely difficult. We previously described the crystal structures of both the inner membrane transporter CusA¹ and the membrane fusion protein CusB² of the CusCBA efflux system^{3,4} of *E. coli*. Here we report the co-crystal structure of the CusBA efflux complex, showing that the transporter (or pump) CusA, which is present as a trimer, interacts with six CusB protomers and that the periplasmic domain of CusA is involved in these interactions. The six CusB molecules seem to form a continuous channel. The affinity of the CusA and CusB interaction was found to be in the micromolar range. Finally, we have predicted a three-dimensional structure for the trimetric CusC outer membrane channel and developed a model of the tripartite efflux assemblage. This CusC₃-CusB₆-CusA₃ model shows a 750-kilodalton efflux complex that spans the entire bacterial cell envelope and exports Cu I and Ag I ions.

In Gram-negative bacteria, efflux systems of the resistance-nodulation-division (RND) family have major roles in intrinsic and acquired tolerance to antibiotics and toxic compounds^{5,6}. They are key components that Gram-negative pathogens use to overcome toxic environments that are unfavourable for their survival. An RND efflux pump^{7–14} works in conjunction with a periplasmic membrane fusion protein^{15–18} and an outer membrane channel to form a functional protein complex^{19,20}. In *E. coli*, one such tripartite efflux system, CusCBA, is responsible for extruding Cu I and Ag I ions, which are biocidal^{3,4}. CusA is a large RND efflux pump that depends on the proton-motive force, is present in the inner membrane and comprises 1,047 amino acids^{3,4}. CusC is a 457 amino acid protein that forms an outer membrane channel^{3,4}. The membrane fusion protein CusB, which consists of 379 amino acids, bridges CusA and CusC to form the tripartite efflux complex CusCBA^{3,4}. This three-component system spans the entire cell envelope of *E. coli* and exports Cu I and Ag I ions. Between the *cusC* and *cusB* genes, there is a small chromosomal gene that encodes a periplasmic protein, CusF⁴. CusF functions as a chaperone carrying Cu I and Ag I ions to CusCBA^{21,22}.

We previously reported the crystal structure of the full-length CusB membrane fusion protein at a resolution of 3.40 Å, revealing four linearly arranged domains (domains 1–4), which constitute approximately 80% of the CusB protein². Overall, CusB is folded into an elongated structure, which is about 120 Å long and about 40 Å wide. The first three domains (domains 1–3) of the protein consist mostly of

β-strands. However, the fourth domain (domain 4) consists entirely of α-helices and is folded into a three-helix bundle structure.

We also previously determined the crystal structure of the full-length CusA RND pump (at 3.52-Å resolution), including approximately 98% of the amino acids¹. The structure suggests that CusA exists as a homotrimer. Each subunit of CusA consists of 12 transmembrane helices (TM1–TM12) and a large periplasmic domain formed by two periplasmic loops, one between TM1 and TM2 and another between TM7 and TM8. The periplasmic domain of CusA can be divided into a pore domain (comprising subdomains PN1, PN2, PC1 and PC2) and a CusC-docking domain (containing subdomains DN and DC). Through the use of lysine–lysine crosslinking and mass spectrometry, it was determined that domain 1 of CusB is in direct contact with subdomain PN1 of the CusA efflux pump².

Here we describe the co-crystal structure of the CusBA heavy-metal efflux complex. We used molecular replacement with single-wavelength anomalous dispersion to determine the structure (Supplementary Table 1 and Supplementary Fig. 1), revealing that the periplasmic domain of each protomer of CusA interacts specifically with two elongated molecules of CusB (molecules 1 and 2) (Fig. 1). The orientation of the CusBA complex is shown in Fig. 1 (and the terms upper and lower refer to this orientation). The two CusB adaptors are tilted at an angle of about 50° with respect to the membrane surface and establish a close fit with the CusA transporter at the concave surface formed by domains 1 and 2 of the adaptor. Molecule 1 of CusB is mainly in contact with the upper regions of PN2 and PC1 and the DN subdomain of CusA. Molecule 2 of CusB, however, predominantly bridges the upper regions of PC1 and PC2, as well as the subdomain DC of the pump. These two adaptor molecules are also seen to contact one another specifically, primarily through domains 1, 2 and 3 of these two elongated molecules. The trimetric CusA pump therefore directly contacts six CusB adaptor molecules, which form a channel at the top of the CusA trimer (Figs 1 and 2).

Intriguingly, molecule 1 of CusB interacts predominantly with CusA through charge–charge interactions. The amino acid residues K95, D386, E388 and R397 of this CusB molecule form four salt bridges, with D155, R771, R777 and E584 of CusA, respectively (Fig. 3a). In addition, T89, the backbone oxygen of N91, and R292 of molecule 1 of CusB form hydrogen bonds with K594, R147 and the backbone oxygen of Q198 of CusA, respectively, to secure the interaction. However, the interaction of molecule 2 of CusB with CusA seems to be governed principally by charge–dipole and dipole–dipole interactions. Specifically, Q108, S109, S253 and N312 of molecule 2 of CusB form hydrogen bonds with Q785, Q194, D800 and Q198 of CusA, respectively (Fig. 3b). The backbone oxygens of L92 and T335 of this CusB molecule also contribute two additional hydrogen bonds with the side chains of K591 and T808 of the CusA pump, anchoring the proteins.

¹Department of Chemistry, Iowa State University, Ames, Iowa 50011, USA. ²Bioinformatics and Computational Biology Interdepartmental Graduate Program, Iowa State University, Ames, Iowa 50011, USA.

³NE-CAT and Department of Chemistry and Chemical Biology, Cornell University, Building 436E, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, USA. ⁴Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, Iowa 50011, USA. ⁵Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA.

*These authors contributed equally to this work.

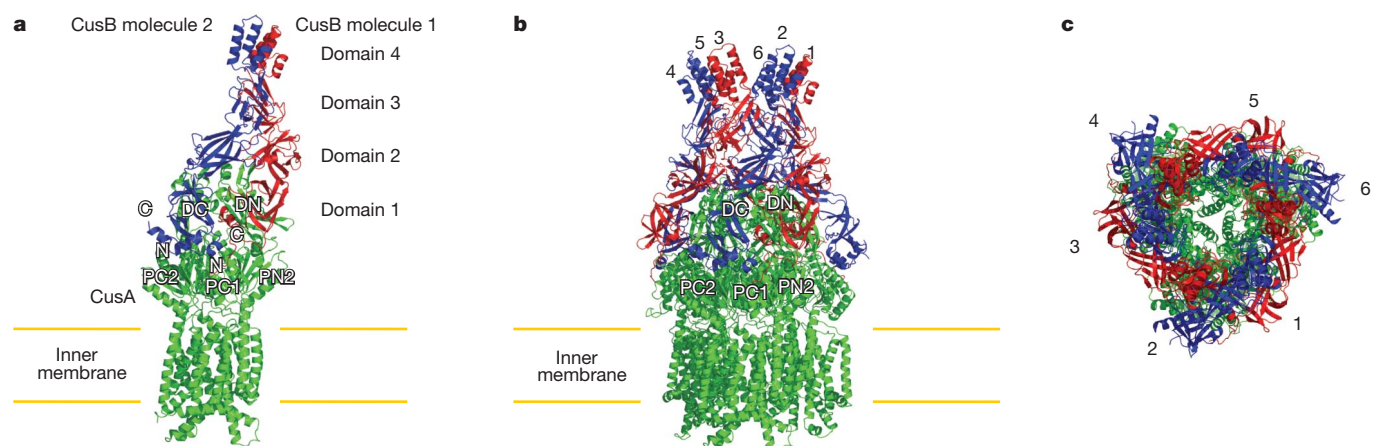


Figure 1 | Structure of the CusBA efflux complex. **a**, Ribbon diagram of the structures of one CusA protomer (green) and two CusB protomers (red and blue) in the asymmetrical unit of the crystal lattice. DC, DN, PC1, PC2 and PN2 are subdomains of CusA. C, carboxy terminus; N, amino terminus. **b**, Side view

of the CusBA efflux complex. Each subunit of CusA is shown in green. Molecules 1, 3 and 5 of CusB are in red. Molecules 2, 4 and 6 of CusB are in blue. **c**, Top view of the CusBA efflux complex. Each subunit of CusA is in green. Molecules 1, 3 and 5 of CusB are in red. Molecules 2, 4 and 6 of CusB are in blue.

In CusB–CusB interactions, molecule 1 of CusB makes a close contact with molecule 2 of CusB. Domains 1–3 of these two molecules are involved in this binding. E118, Y119, R186, E252 and R292 of molecule 1 of CusB participate to form hydrogen bonds with T139, D142, T206, N312 and N113 of molecule 2 of CusB, respectively (Fig. 4a). Furthermore, molecule 1 of CusB also contributes in contacting molecule 6 of CusB, which is anchored to the next subunit of CusA (Fig. 4b). Most of the interactions are between domains 2 and 3 of these two molecules. In particular, N113, N228 and N312 of molecule 1 of CusB pair with R292, the backbone oxygen of A126, and E252 of molecule 6 of CusB, respectively, to form three hydrogen bonds. In addition, D142 of molecule 1 of CusB participates to form two hydrogen bonds, with Y119 and R297 of molecule 6 of CusB, securing the binding.

The side-by-side packing arrangement of six CusB protomers allows a funnel-like structure to be created, with the central channel being formed along the crystallographic three-fold symmetry axis (Fig. 2). Domain 1 and the lower half of domain 2 of CusB primarily create a cap-like structure, whereas the upper half of domain 2, together with domains 3 and 4, contributes to the central channel of the funnel. The inner surface of the cap fits closely with the outer surface of the upper

portion of the periplasmic domain of the CusA trimer. The channel formed above the cap of the adaptor is ~ 62 Å in length with an average internal diameter of ~ 37 Å. Thus, the interior of the channel gives rise to a large elongated cavity with a volume of $\sim 65,000$ Å³. The lower half of the channel is primarily created by β -barrels, whereas the upper half is an entirely α -helical tunnel. The diameter of the channel is gradually constricted and then dilates as it approaches the outer membrane. Thus, the α -helices of domain 4 create an inverted conical structure. The narrowest section of the central channel is located at residue D232 of each CusB protomer, which is close to the hinge region between domains 3 and 4 of CusB. The widest section of the channel appears to form at the top edge, with an inner diameter of ~ 56 Å. The inner surface of the channel is predominantly negatively charged, as indicated by the electrostatic surface diagram (Fig. 5), suggesting that the interior surface of the channel may have the capacity to bind to positively charged metal ions.

We used isothermal titration calorimetry to calculate the binding affinity of CusB to the CusA pump. The data indicate an equilibrium dissociation constant of 5.1 ± 0.3 μ M (Supplementary Fig. 2).

Previously, it was found that the amino- and carboxy-terminal ends (residues 29–88 and 386–407) of the CusB adaptor are intrinsically

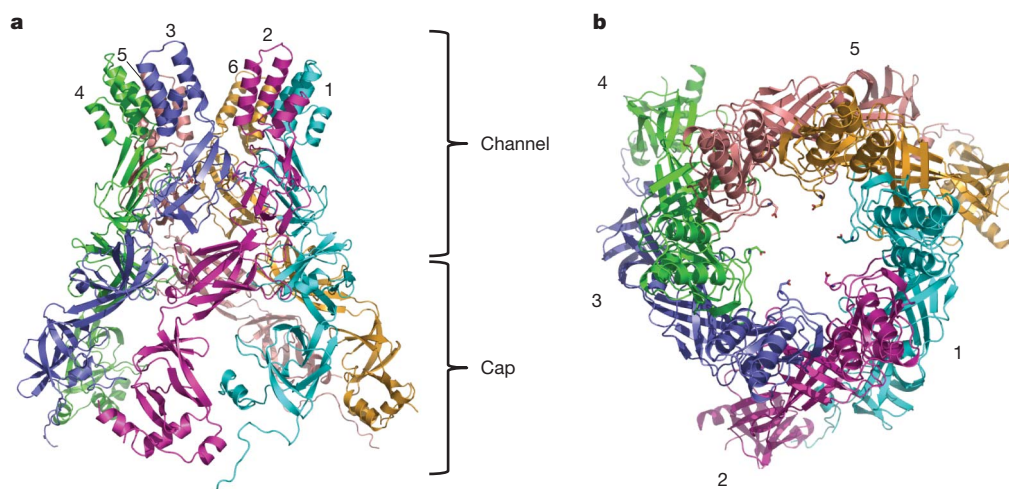


Figure 2 | Structure of the hexameric CusB channel. **a**, Side view of the hexameric CusB channel. The six molecules of CusB are shown as ribbons (cyan, molecule 1; dark pink, molecule 2; purple, molecule 3; green, molecule 4; pale pink, molecule 5; and orange, molecule 6). Domain 1 and the lower half of domain 2 of the molecules of the CusB hexamer form a cap, whereas the upper

half of domain 2, together with domains 3 and 4, of these molecules creates a channel. **b**, Top view of the hexameric CusB channel. The six molecules of CusB are depicted as in **a**. The D232 residues from the six subunits of CusB, which together form the narrowest region of the central channel, are depicted in stick format. The internal diameter of this region is ~ 18 Å.

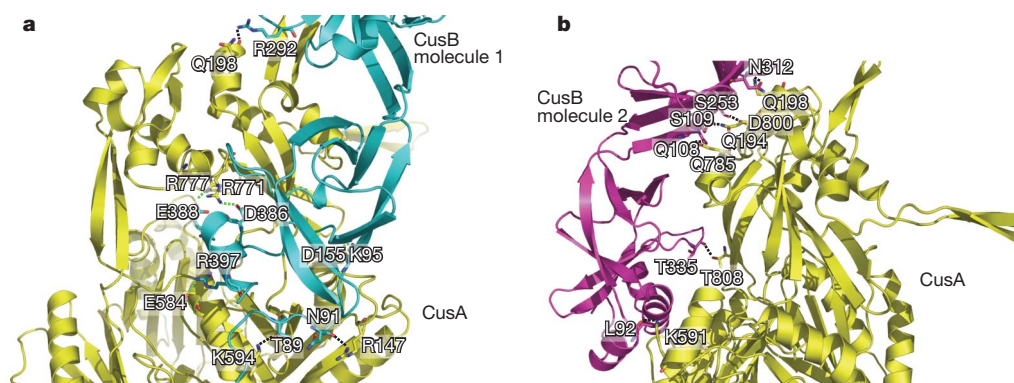


Figure 3 | CusA–CusB interactions. CusB and CusA are depicted as ribbons, with the residues that make important interactions depicted in stick format (red, oxygen; blue, nitrogen). CusB ribbon colour matches that in Fig. 2. **a**, The interactions between CusA and molecule 1 of CusB. In this CusB molecule, residues K95, D386, E388 and R397 form four salt bridges (green) with D155, R771, R777 and E584 of CusA, respectively. In addition, T89, the backbone oxygen of N91, and R292 of molecule 1 of CusB form hydrogen bonds with

K594, R147, and the backbone oxygen of Q198 of CusA. **b**, The interactions between CusA and molecule 2 of CusB. In molecule 2 of CusB, residues Q108, S109, S253 and N312 form hydrogen bonds with Q785, Q194, D800 and Q198 of CusA, respectively. The backbone oxygens of L92 and T335 of this CusB molecule also contribute two additional hydrogen bonds, with the side chains of K591 and T808 of CusA, anchoring the proteins.

disordered and cannot be identified in the electron density maps of CusB crystals². Here the co-crystal structure suggests that these regions form several short α -helices. In molecule 1 of CusB, residues 392–399 at the C-terminal end form a short α -helix. However, residues 79–95 of the N terminus feature a long random coil, and these amino acids are located immediately outside the cleft that is formed between subdomains PC1 and PC2 of the CusA pump. For molecule 2 of CusB, the N-terminal residues 79–85 and 86–92 participate to form a random coil and a short α -helix, respectively. However, the C-terminal residues 382–392 and 394–400 appear to create two short α -helices. Like molecule 1 of CusB, the N terminus of molecule 2 of CusB is near the periplasmic cleft of the pump (Fig. 1).

It has been proposed that the N-terminal residues M49, M64 and M66 of CusB form a three-methionine metal-binding site²³. Although these three methionine residues cannot be identified in the electron density maps of our co-crystal, the co-crystal structure shows that the N-terminal tails of both molecules 1 and 2 of CusB are located outside the cleft formed between PC1 and PC2 of the CusA pump (Fig. 1). Thus, it is possible that CusB might help to transfer the metal ions by way of the N-terminal three-methionine binding site into the periplasmic cleft of CusA. Indeed, a similar suggestion has been made for the

adaptor protein AcrA, which is also a component of a bacterial RND efflux pump, although this protein might assist in transporting a drug from the periplasm into the pump²⁴.

The co-crystal structure presented here highlights the structural importance of the periplasmic membrane fusion protein. Given that six CusB molecules assemble to form a channel, the adaptor CusB is likely to be involved in the active extrusion of metal ions.

We next constructed a CusCBA model based on the CusBA crystal complex structure and the predicted CusC model. The final structural model of CusC₃–CusB₆–CusA₃ is a 750-kDa tripartite efflux complex that spans both the inner and outer membrane of *E. coli* and extrudes Cu I and Ag I ions (Supplementary Fig. 3).

We believe that CusA can take up metal ions from both the periplasm and cytoplasm, using the methionine-residue ion relay network¹. Metal ions could enter the three-methionine binding site of CusA, formed by M573, M623 and M672, which is inside the cleft between subdomains PN2 and PC1 on the periplasmic portion of CusA, or these ions could enter by way of the methionine pairs within the transmembrane domain of the CusA pump. It has been demonstrated that the chaperone CusF can directly transfer its bound Cu I to the CusB membrane fusion protein²⁵. Thus, it is likely that CusF is responsible for delivering

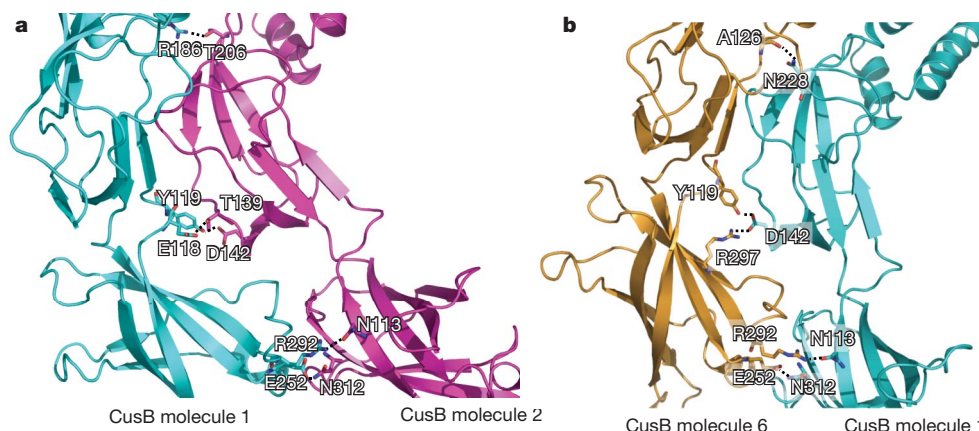


Figure 4 | CusB–CusB interactions. CusB molecules are depicted as ribbons (colour matched to Fig. 2), with the residues that make important interactions depicted in stick format (red, oxygen; blue, nitrogen). **a**, The interactions between molecules 1 and 2 of CusB. Residues E118, Y119, R186, E252 and R292 of molecule 1 of CusB participate to form hydrogen bonds with residues T139, D142, T206, N312 and N113 of molecule 2 of CusB, respectively. These hydrogen-bonded distances are 2.7 Å, 2.7 Å, 3.0 Å, 3.0 Å and 3.0 Å, respectively.

b, The interactions between molecules 1 and 6 of CusB. Residues N113, N228 and N312 of molecule 1 of CusB pair with R292, the backbone oxygen of A126, and E252 of molecule 6 of CusB, respectively, to form three hydrogen bonds. D142 of molecule 1 of CusB also participates to form two hydrogen bonds, with Y119 and R297 of molecule 6 of CusB. These hydrogen-bonded distances are 2.7 Å, 2.8 Å, 3.1 Å, 2.7 Å and 2.8 Å, respectively.

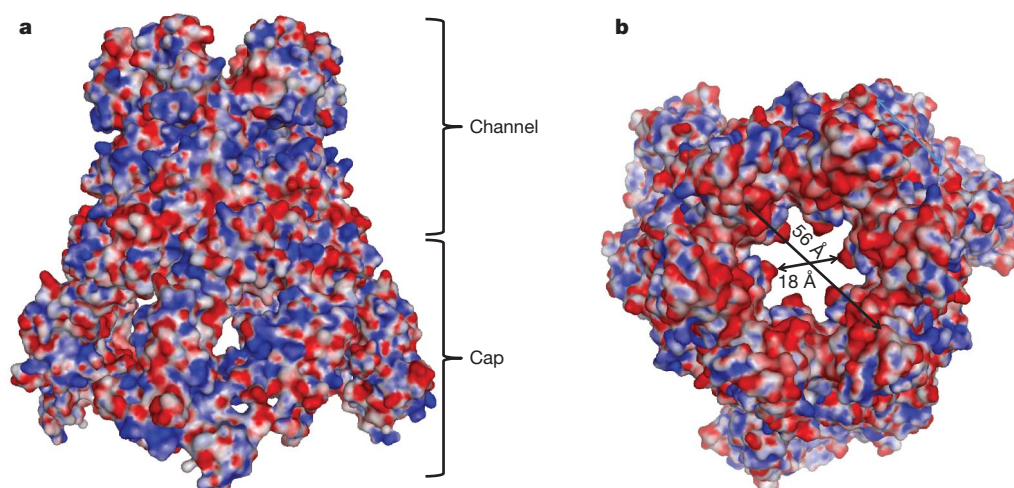


Figure 5 | Electrostatic surface potential of CusB. Electrostatic surface potential of hexameric CusB. Blue ($> 15 k_B T$) and red ($< -15 k_B T$) indicate the positively and negatively charged areas, respectively, of the protein, where k_B is the Boltzmann constant and T is absolute temperature. White denotes areas

between $-15 k_B T$ and $15 k_B T$. **a**, Side view. This view shows the cap and channel regions formed by the CusB hexamer. **b**, Top view. The widest section of the hexameric channel appears to form at the top edge, with an internal diameter of ~ 56 Å, and the narrowest section of the channel has a diameter of 18 Å.

metal ions to the CusCBA tripartite efflux system in the periplasm. There is a possibility that the first step of metal transport through the periplasmic cleft of the CusA pump involves direct transfer from CusF to the previously proposed three-methionine metal-binding site²³ (M49, M64 and M66) at the long N-terminal tail of CusB, situated near the cleft between PN2 and PC1 of the CusA pump. Thus, the second step could be the delivery of the bound metal ion from CusB to the three-methionine cluster (M573, M623 and M672) inside the periplasmic cleft of CusA¹. The bound metal ion could then be released to the nearest methionine pair (M271–M755), which is located directly above the three-methionine metal-binding site¹, from which the ion could be released from CusA and eventually reach the CusC channel for final extrusion. It is not yet known whether the interior of the hexameric CusB channel forms part of the extrusion pathway for Cu I and Ag I ions. The exact mechanism by which this tripartite efflux system operates must await confirmation through elucidating additional crystal structures of the CusCBA tripartite complex.

METHODS SUMMARY

Crystallization of CusBA. The procedures for cloning, expression and purification of the CusA and CusB proteins have been described previously^{1,2}. Co-crystals of the CusBA complex were obtained using sitting-drop vapour diffusion. A 2 μ l protein solution containing 0.1 mM CusA and 0.1 mM CusB in buffer solution containing 20 mM Na-HEPES, pH 7.5, and 0.05% (w/v) CYMAL-6 was mixed with 2 μ l of reservoir solution containing 10% PEG 6000, 0.1 M Na-HEPES, pH 7.5, 0.1 M ammonium acetate and 20% glycerol. The resultant mixture was equilibrated against 500 μ l of the reservoir solution. The co-crystallization conditions for CusA (native)–CusB (selenomethionine-substituted, SeMet) were the same as those for the native CusBA complex. Co-crystals of CusBA grew to full size in the drops within 2 months. Cryoprotection was achieved by raising the glycerol concentration stepwise to 30% in 5% increments.

Structural determination and refinement. Data for the CusA (native)–CusB (SeMet) co-crystal were collected (Supplementary Table 1). Single-wavelength anomalous dispersion (SAD) phasing using the program Phaser²⁶ was used to obtain experimental phases in addition to the phases from the structural model of the apo form of CusA. Phases were subjected to density modification and phase extension to 2.90-Å resolution using the program RESOLVE²⁷. The full-length CusB protein contains nine methionine residues, and six selenium sites per CusB molecule (12 selenium sites per asymmetrical unit) were identified. After tracing the initial model manually using the program Coot²⁸, the model was refined against the native data at 2.90-Å resolution using translation, libration and screw-rotation (TLS) refinement adopting a single TLS body as implemented in the program PHENIX²⁹, leaving 5% of the reflections in the free R set. Iterations of refinement using PHENIX²⁹ and the software suite CNS³⁰ and model building in

Coot²⁸ led to the current model, which consists of 1,686 amino acid residues (residues 4–1,043 of CusA, residues 79–400 of molecule 1 of CusB and residues 79–402 of molecule 2 of CusB).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 June; accepted 6 December 2010.

- Long, F. *et al.* Crystal structures of the CusA heavy-metal efflux pump suggest methionine-mediated metal transport mechanism. *Nature* **467**, 484–488 (2010).
- Su, C.-C. *et al.* Crystal structure of the membrane fusion protein CusB from *Escherichia coli*. *J. Mol. Biol.* **393**, 342–355 (2009).
- Franke, S., Grass, G. & Nies, D. H. The product of the *ybdE* gene of the *Escherichia coli* chromosome is involved in detoxification of silver ions. *Microbiology* **147**, 965–972 (2001).
- Franke, S., Grass, G., Rensing, C. & Nies, D. H. Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli*. *J. Bacteriol.* **185**, 3804–3812 (2003).
- Tseng, T. T. *et al.* The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J. Mol. Microbiol. Biotechnol.* **1**, 107–125 (1999).
- Nies, D. H. Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol. Rev.* **27**, 313–339 (2003).
- Murakami, S., Nakashima, R., Yamashita, E. & Yamaguchi, A. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* **419**, 587–593 (2002).
- Yu, E. W., McDermott, G., Zgurskaya, H. I., Nikaido, H. & Koshland, D. E. Jr. Structural basis of multiple drug binding capacity of the AcrB multidrug efflux pump. *Science* **300**, 976–980 (2003).
- Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T. & Yamaguchi, A. Crystal structure of a multidrug transporter reveal a functionally rotating mechanism. *Nature* **443**, 173–179 (2006).
- Seeger, M. A. *et al.* Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science* **313**, 1295–1298 (2006).
- Sennhauser, G., Amstutz, P., Briand, C., Storcheggger, O. & Grütter, M. G. Drug export pathway of multidrug exporter AcrB revealed by DARPin inhibitors. *PLoS Biol.* **5**, e7 (2007).
- Yu, E. W., Aires, J. R., McDermott, G. & Nikaido, H. A periplasmic-drug binding site of the AcrB multidrug efflux pump: a crystallographic and site-directed mutagenesis study. *J. Bacteriol.* **187**, 6804–6815 (2005).
- Sennhauser, G., Bukowska, M. A., Briand, C. & Grütter, M. G. Crystal structure of the multidrug exporter MexB from *Pseudomonas aeruginosa*. *J. Mol. Biol.* **389**, 134–145 (2009).
- Su, C.-C. *et al.* Conformation of the AcrB multidrug efflux pump in mutants of the putative proton relay pathway. *J. Bacteriol.* **188**, 7290–7296 (2006).
- Higgins, M. K., Bokma, E., Koronakis, E., Hughes, C. & Koronakis, V. Structure of the periplasmic component of a bacterial drug efflux pump. *Proc. Natl Acad. Sci. USA* **101**, 9994–9999 (2004).
- Akama, H. *et al.* Crystal structure of the membrane fusion protein, MexA, of the multidrug transporter in *Pseudomonas aeruginosa*. *J. Biol. Chem.* **279**, 25939–25942 (2004).
- Mikolasko, J., Bobyk, K., Zgurskaya, H. I. & Ghosh, P. Conformational flexibility in the multidrug efflux system protein AcrA. *Structure* **14**, 577–587 (2006).

18. Symmons, M., Bokma, E., Koronakis, E., Hughes, C. & Koronakis, V. The assembled structure of a complete tripartite bacterial multidrug efflux pump. *Proc. Natl Acad. Sci. USA* **106**, 7173–7178 (2009).
19. Koronakis, V., Sharff, A., Koronakis, E., Luisi, B. & Hughes, C. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* **405**, 914–919 (2000).
20. Akama, H. *et al.* Crystal structure of the drug discharge outer membrane protein, OprM, of *Pseudomonas aeruginosa*. *J. Biol. Chem.* **279**, 52816–52819 (2004).
21. Xue, Y. *et al.* Cu^I recognition via cation- π and methionine interactions in CusF. *Nature Chem. Biol.* **4**, 107–109 (2008).
22. Loftin, I. R., Franke, S., Blackburn, N. J. & McEvoy, M. M. Unusual Cu^I/Ag^I coordination of *Escherichia coli* CusF as revealed by atomic resolution crystallography and X-ray absorption spectroscopy. *Protein Sci.* **16**, 2287–2293 (2007).
23. Bagai, I., Liu, W., Rensing, C., Blackburn, N. & McEvoy, M. M. Substrate-linked conformational change in the periplasmic component of Cu^I/Ag^I efflux system. *J. Biol. Chem.* **282**, 35695–35702 (2007).
24. Gerken, H. & Misra, R. Genetic evidence for functional interactions between TolC and AcrA proteins of a major antibiotic efflux pump of *Escherichia coli*. *Mol. Microbiol.* **54**, 620–631 (2004).
25. Bagai, I., Rensing, C., Blackburn, N. J. & McEvoy, M. M. Direct metal transfer between periplasmic proteins identifies a bacterial copper chaperone. *Biochemistry* **47**, 11408–11414 (2008).
26. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
27. Terwilliger, T. C. Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Crystallogr. D* **57**, 1755–1762 (2001).
28. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
29. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58**, 1948–1954 (2002).
30. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work is supported by National Institutes of Health grants R01GM074027 (E.W.Y.), R01GM086431 (E.W.Y.), R01GM081680 (R.L.J.) and R01GM072014 (R.L.J.). This work is based on research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source is supported by the US Department of Energy, Office of Basic Energy Sciences, under contract DE-AC02-06CH11357.

Author Contributions C.-C.S., F.L. and E.W.Y. designed the research. C.-C.S. and F.L. performed experiments. M.T.Z. and R.L.J. performed the docking. C.-C.S., F.L., K.R.R. and E.W.Y. performed the model building and refinement. C.-C.S., F.L., R.L.J. and E.W.Y. wrote the paper.

Author Information Atomic coordinates and structure factors for the CusBA structure have been deposited with the Protein Data Bank (<http://www.pdb.org>) under accession code 3NE5. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to E.W.Y. (ewyu@iastate.edu).

METHODS

Crystallization of CusBA. The procedures for cloning, expression and purification of the CusA and CusB proteins have been described previously^{1,2}. Co-crystals of the CusBA complex were obtained using sitting-drop vapour diffusion. A 2 μ l protein solution containing 0.1 mM CusA and 0.1 mM CusB in buffer solution containing 20 mM Na-HEPES, pH 7.5, and 0.05% (w/v) CYMAL-6 was mixed with 2 μ l of reservoir solution containing 10% PEG 6000, 0.1 M Na-HEPES, pH 7.5, 0.1 M ammonium acetate and 20% glycerol. The resultant mixture was equilibrated against 500 μ l of the reservoir solution. The co-crystallization conditions for CusA (native)–CusB (selenomethionine-substituted, SeMet) were the same as those for the native CusBA complex. Co-crystals of CusBA grew to full size in the drops within 2 months. Cryoprotection was achieved by raising the glycerol concentration stepwise to 30% in 5% increments.

Structural determination and refinement. Data for the CusA (native)–CusB (SeMet) co-crystal were collected (Supplementary Table 1). Single-wavelength anomalous dispersion (SAD) phasing using the program Phaser²⁶ was used to obtain experimental phases in addition to the phases from the structural model of the apo form of CusA. Phases were subjected to density modification and phase extension to 2.90-Å resolution using the program RESOLVE²⁷. The full-length CusB protein contains nine methionine residues, and six selenium sites per CusB molecule (12 selenium sites per asymmetrical unit) were identified. After tracing the initial model manually using the program Coot²⁸, the model was refined against the native data at 2.90-Å resolution using translation, libration and screw-rotation (TLS) refinement adopting a single TLS body as implemented in the program PHENIX²⁹, leaving 5% of the reflections in the free *R* set. Iterations of refinement using PHENIX²⁹ and the software suite CNS³⁰ and model building in Coot²⁸ led to the current model, which consists of 1,686 amino acid residues (residues 4–1,043 of CusA, residues 79–400 of molecule 1 of CusB and residues 79–402 of molecule 2 of CusB).

Isothermal titration calorimetry. We used isothermal titration calorimetry to examine the binding of the purified CusB adaptor to the purified CusA pump. Measurements were performed on a VP microcalorimeter (MicroCal) at 25 °C. Before titration, the CusA and CusB proteins were thoroughly dialysed against buffer containing 20 mM Na-HEPES, pH 7.5, and 0.05% CYMAL-6, respectively. Protein concentrations were determined using the Bradford assay. The CusA protein sample was then adjusted to a final monomeric concentration of 14 μ M. The CusB protein solution, consisting of 350 μ M monomeric CusB in 20 mM Na-HEPES, pH 7.5, and 0.05% CYMAL-6, was prepared as the titrant. The samples were degassed before they were loaded into the cell and syringe. Binding experiments were carried out with the CusA protein solution (1.5 ml) in the cell and the

CusB protein solution as the injectant. Ten microlitre injections of the CusB solution were used for data collection.

Injectations occurred at intervals of 300 s, and the duration time of each injection was 10 s. Heat transfer (μ cal s^{−1}) was measured as a function of elapsed time (s). The mean enthalpies measured from injection of the ligand in the buffer were subtracted from raw titration data before data analysis with ORIGIN software (MicroCal). Titration curves were fitted by a nonlinear least squares method to a function for the binding of a ligand to a macromolecule. Nonlinear regression fitting to the binding isotherm provided us with the equilibrium binding constant ($K_a = 1/K_d$, where K_a is the association constant and K_d is the dissociation constant) and enthalpy of binding (ΔH). Based on the values of K_a , the change in free energy (ΔG) and entropy (ΔS) were calculated with the equation $\Delta G = -RT \times \ln K_a = \Delta H - T\Delta S$, where R is 1.9872 cal K^{−1} mol^{−1} and T is 273 K. Calorimetry trials were also carried out in the absence of CusA in the same experimental conditions. No change in heat was observed in the injections throughout the experiment.

Structural prediction and docking of CusC onto CusBA. A homology model of CusC was generated using the crystal structure of OprM (Protein Data Bank accession code 1WP1)²⁰ as a template through the I-TASSER server³¹. Alignment of protein sequences suggests that these two channel-forming proteins have 42% identity. A steered molecular dynamics simulation was then used to rigidly dock the trimeric CusC channel to the crystal form of the hexameric CusB channel bound to the trimeric CusA pump. During the rigid-body docking process, the molecules were aligned with the *z* axis. CusB was then held rigidly, allowing CusC to pull along the *z* axis towards CusB until no further change in the centre of mass was measured.

We next investigated the possible interaction of CusC with domain 2 of CusB, which also brings it into close proximity to CusA. First, the trimeric CusA pump and domains 1 and 2 of the hexameric CusB channel were held rigidly, while domains 3 and 4 of each CusB monomer were pulled outwards using customized harmonic forces to open the annulus of the hexameric channel. Next, we manually inserted CusC into the opened CusB ring. Finally, simulated annealing followed by energy minimization was performed to relax CusB and CusC from these positions and allow them to assume a lower energy conformation. All simulations were performed using NAMD³² and the CHARMM27+CMAP force field³³.

31. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5**, 725–738 (2010).
32. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
33. Feller, S. E. & MacKerell, A. D. Jr. An improved empirical potential energy for molecular simulations of phospholipids. *J. Phys. Chem. B* **104**, 7510–7515 (2000).

CAREERS

POSTDOCS NIH postdoc stipend could rise 4% under Obama's proposal **p.565**

COLUMN Is an overseas postdoc crucial for career progression? **p.565**

NATUREJOBS For the latest career listings and advice www.naturejobs.com

IMAGEZOO/CORBIS



CAREER CHOICES

The mobility imperative

Changing labs is crucial for the next generation of European scientists. But mobility can be a mixed blessing.

BY QUIRIN SCHIERMEIER

Marine biologist Martin Pfannkuchen has been on the move for much of his young career. He realized early on that Swabia, the southwest region of Germany he came from, was not an ideal base for his work on the cell biology of marine

sponges and algae. With two national borders and the European Alps between him and the nearest coast, he had limited access to water or a coastal lab. Already, as a research associate in the early stage of his career, he has had valuable experiences abroad. During his PhD at the University of Stuttgart in Germany, Pfannkuchen spent a month at the

Natural History Museum in London. After his PhD, he considered continuing his research abroad at the University of Hawaii or Kyoto University in Japan, where he had been offered postdoctoral positions.

But wishing to be close to the sea and to his Croatian partner, nine months after finishing his PhD he resettled in Rovinj, a town on Croatia's Adriatic coast that is home to a small marine research centre at the Ruđer Bošković Institute. Although not a major scientific player, Croatia has all that Pfannkuchen needs. "One has to know how to work self-sufficiently here, but the equipment and grant opportunities are fine, and my faculty prepared me well for this situation," he says. "The best thing is that I can work at sea any day and I have fresh samples all the time."

MOTIVATING MANTRA

Changing countries has become a rite of passage for many young researchers, especially in Europe, where cross-border mobility is common. The call for mobility has become the motivating mantra of organizations such as the Marie Curie fellowship programme, which promotes and supports mobility across Europe. In Germany, for example, to avoid academic 'inbreeding', in which universities hire their own graduates as professors, university tenure rules require scientists to change labs during the course of their postdoc or graduate education, and trips abroad to the United States or elsewhere are all but expected. In many countries, recruiters and funding agencies see international mobility as a mark of an applicant's ability and dedication, making changing labs a key to scientists' professional success almost everywhere.

Nevertheless, the practical professional outcome of mobility is hard to pin down — raising suspicions that it is sometimes undertaken for its own sake, as a means to boost a CV rather than a way to expand one's knowledge base. Crossing borders is often a fruitful enterprise — an opportunity to experience different lab cultures, acquire new skills, learn new methods and establish personal contacts and networks that can be important for future career progressions. But moving for the sake of moving may have little effect on one's capability or marketability. And there are other implications: for later-career postdocs and professors, pensions could be affected, for example. Graduate students and others should therefore consider their own long-term costs and benefits. Although supervisors may ►

► provide guidance, they also have their own research agenda and priorities.

Mobility numbers vary widely depending on region. At leading research universities such as the University of Cambridge, UK, more than 40% of researchers are from abroad. Across Europe as a whole, however, no more than about 7% of research-and-development personnel work outside their native country, according to figures compiled in 2007 by

the European Commission (EC) Joint Research Centre in Seville, Spain.

Regardless, mobility is integral to the EC's scientific workforce initiatives. The European Research Area Board, a consultative body that advises the commission, has set a target of 20% of European Union (EU) doctoral candidates working outside their home countries by 2020 — roughly a tripling of current figures. Outdated tenure, pension

and social-security systems all hamper the scientific mobility of EU researchers, the group said in a report released in 2009.

And mobility issues will be freshly addressed as the EU moves towards its 2014–20 Framework programme for research, the priorities and budget of which are being negotiated at present. Under discussion are a European pension fund for mobile researchers (see *Nature* 467, 489–491; 2010), and the establishment of doctoral programmes that are specific for non-national PhD students at several universities across the EU. “Start encouraging mobility at the doctoral level and you’ll get a more international research environment in Europe,” says John Smith, deputy secretary-general of the European University Association (see Column, page 565).

In addition, 47 European countries participate in the Bologna process, which was launched in 1999 and aims to improve mobility by uniting education systems throughout the continent with common principles for training and the supervision of students, and reducing the time to a degree. All this suggests that, for fledgling EU scientists, the pressure to be mobile may become even stronger in the future.

ALL THE RIGHT MOVES

Stints abroad do have their benefits. “The best lab for the type of research you’re doing is normally not around the corner — it’s in fact very likely to be somewhere else,” says António Coutinho, director of the Gulbenkian Science Institute in Oeiras, Portugal, which

runs three international PhD programmes in the life sciences.

That has been the experience of Martin Pešl, a PhD student at the International Clinical Research Center Brno in the Czech Republic. “I had the chance to work with some of the leading researchers in my field — that was just fantastic,” he says of his six-month stay at the Mayo Clinic in Rochester, Minnesota, from where he returned in December. Pešl, who studies the use of human embryonic stem cells to treat cardiovascular diseases, plans to return to the Mayo Clinic next year to do postdoctoral research, but hopes eventually to start an independent research group in the Czech Republic. “I know where my home is and I love my family,” he says.

THE LANGUAGE BARRIER

Although career advisers advocate early-career mobility, some admit that there are downsides. Graduate students considering a period abroad should consider the effect of cultural differences, in an attempt to avoid stress and conflict. Those moving from Europe to places such as Japan (and vice versa) may benefit from cultural coaching to better understand unfamiliar hierarchies, conflict management and body language. And, of course, language is an issue. Working in a lab without knowing the local language may be feasible, but this is more difficult in the case of everyday life. “Let’s face it, some countries — especially those where there is little incoming mobility — embrace diversity less easily than others,” says András Dinnyés, a biomedical scientist who did postdoctoral research in Japan and China before starting the biotech company BioTalentum in Gödöllő, Hungary.

Students should seek advice, but be aware of its source. “Don’t be overly dependent on your supervisor,” says Karen Vandeveld, a research-policy adviser at Ghent University in Belgium, who studies mobility among Flemish researchers. “PhD students are in a vulnerable position, especially if their supervisors are not so good. There are lots of reasons why some might not encourage mobility, for example, because they don’t want to lose their students.”

Empirical data support the idea that the ‘mobility imperative’ restricts women in science and creates obstacles for academics, says Louise Ackers, chair in European Socio-Legal Studies at the University of Liverpool, UK, who has conducted several research studies on the relationship between mobility and internationalization. Vandeveld says that recruiters often unfairly think that researchers who are not mobile because of family restrictions are not passionate enough about research.

But Ackers — herself a mother of four — has also found that scientific mobility has evolved, as short trips to conferences and collaborators, virtual meetings and remote access to research data can reduce the migration of scientists to some extent. “In fields dominated by large

international teams, such as particle physics, successful research no longer demands being where it’s at all the time,” she says. “You may have to be prepared to travel a lot, but not necessarily to live abroad.”

For graduate students, finding funding to move can be a challenge, but there are some mechanisms in place to help. The EC-funded Erasmus and Erasmus Mundus programmes (for student mobility within and outside the EU, respectively) and the Marie Curie fellowships allow thousands of postdocs and PhD students to gain experience abroad every year.

Even so, money is still an issue. A survey of almost 9,000 PhD students throughout Europe in 2009 by the European Council of Doctoral Candidates and Junior Researchers found that more than 20% used personal savings to finance stays abroad or relied on support from partners and relatives. Less than 9% of graduate students were aware of the European Charter for Researchers or the Code of Conduct for the Recruitment of Researchers — recommendations issued by the EC in 2005 for employment conditions and social-security coverage for researchers, including PhD students. These standards addressed issues such as treating PhD candidates as professional scientists, not only as students, and ensuring that they have proper contracts. “It’s a disgrace that

the charter and code make so little difference,” says Snežana Krstić, a chemical engineer who volunteered her time to help conduct the survey. “Many national, and even some EU, institutions haven’t yet implemented the provisions.”

No doubt, crossing borders will increasingly become a part of science, but the career import of mobility remains unclear. “There’s no easy answer to whether researchers are excellent because they are mobile, or whether they are mobile because they



“The best lab for the type of research you’re doing is normally not around the corner — it’s very likely to be somewhere else.”

António Coutinho

are excellent,” says Vandeveld. Nevertheless, young supervisor Pfannkuchen is encouraging his Croatian students to explore international opportunities. “They think that in European labs everything works perfectly, money flows like water, and supervisors let them do whatever they want,” he says. “Alas, many return a bit disenchanted.” ■

Quirin Schiermeier is *Nature’s* German correspondent.



“The best thing is that I can work at sea any day and I have fresh samples all the time.”

Martin Pfannkuchen

COLUMN

Mobility rethought

Lars H. Breimer, Michael E. Breimer and Douwe D. Breimer say doing a postdoc abroad is unnecessary.

IMAGES.COM/CORBIS

It used to be almost an act of faith that a researcher should undertake a postdoctoral position abroad. In the days before cheap air travel and the Internet, this was the only way to gain international experience and exchange ideas with researchers in other countries on a daily basis. It was a once-in-a-lifetime experience. Now, it has become commonplace — and this is not necessarily a change for the better.

We propose that mobility should no longer be about postdocs spending one to three years abroad, but rather about institutions becoming more international by recruiting undergraduates from abroad, facilitating the movement of master's and PhD students from one institution to another and recruiting early-career professors for teaching and research positions.

This is happening in Europe, where universities have made an effort to become more cosmopolitan. Nations are harmonizing programmes to ease movement — for example, by letting students with bachelor's degrees from one country do a master's in another.

PhD positions are increasingly filled by international students. In the sciences, foreign citizens received 46% of the PhDs awarded in the United States in 2007–08, and 40% in the United Kingdom and Switzerland. In Sweden, 33% of those starting a PhD in the same time period were from elsewhere. Universities are also making their undergraduate programmes more international, by setting up foreign campuses and recruiting foreign students to the home campus. Some people take a year off after high school to travel, work or study abroad.

Thus, many researchers have had exposure to foreign countries before reaching the postdoc level. There is no longer the need for international experience during the postdoc that there may once have been. Perhaps the requirement is a myth, kept alive because grant and appointment committee members travelled themselves. Young scientists thinking of going abroad must make sure that any move is in the interests not only of enriching their lives, but also, more importantly, of expanding their CVs.

The postdoc is a key period in the development of a research career. The environment must allow mutual development, so that the visitor does not simply provide technical expertise to the host lab. Postdocs must ensure that their time abroad is worth more than time spent



at home. Ideally, a postdoc is done at an institution with a good reputation in the field, but this need not be in a foreign country. The urge to see the world can be satisfied in other ways.

Most would contend that those who move about in academia fare better than those who do not, but our experiences suggest otherwise. One of us (MEB) had a PhD supervisor who only ever worked at one university. His career did not suffer — he flourished, and all of his PhD students became professors. Most of them did postdocs abroad; but the one who was most successful (not MEB) never left the town, yet rose to be rector of his alma mater. Another of us (DDB) moved only 200 kilometres within the Netherlands: from Groningen to Nijmegen and finally to Leiden, where he supervised 50 PhD students and became rector of the university. The third (LHB) spent 26 years in England and the United States as a lecturer in molecular biology and epidemiology, and in drug development in the pharmaceutical industry, before returning home to work at a teaching hospital.

Once a connection has been made, worldwide collaboration can be fostered in many ways. Before crossing borders, postdocs should consider the career implications, good and bad. ■

Lars H. Breimer is a professor of laboratory medicine at Örebro University Hospital, Sweden. **Michael E. Breimer** is a professor of surgery at Sahlgrenska Academy at the University of Gothenburg, Sweden. **Douwe D. Breimer** is the former rector of Leiden University, the Netherlands. (LHB and MEB are brothers; DDB is not related.)

BIOTECHNOLOGY

Tax credits help firms

A federal funding scheme helped small US biotechnology companies to create and keep jobs in 2010, says a survey. Congress gave US\$1 billion to the Qualifying Therapeutic Discovery Project awards in 2009; eligible firms got up to \$5 million in tax credits and grants. A poll run by Penn Schoen Berland, a market-research firm based in New York, and sponsored by the Biotechnology Industry Organization (BIO) in Washington DC, found that the funding helped 226 companies to create about 6 jobs each, and save 7. The awards helped 80% of respondents to survive the economic downturn. Some 29% had been asked to move abroad, but 59% of those said the awards would keep them at home. James Greenwood, head of BIO, says these data will help efforts to extend the scheme.

POSTDOCS

Stipend raise proposed

In his 2012 budget, US President Barack Obama called for a 4% increase in the National Research Service Award (NRSA) postdoc stipends granted by the Bethesda, Maryland-based National Institutes of Health. The request is just the start of debate with Congress, so the gain isn't assured. In 2010, after two years without change, stipends were raised by 1% to US\$37,740 for first-year postdocs — but the National Postdoctoral Association (NPA) has called for more. A National Academy of Sciences report in 2000 recommended that first-years earn \$45,000. In a letter to Obama on 25 January, the NPA called stipends “unacceptably low”, noting that many US institutions use the NRSA stipend scale.

BIOMEDICINE

Stem-cell opportunities

A stem-cell research centre at the University of California, San Francisco (UCSF), will hire up to 18 biologists, neurologists, immunologists and researchers, and 12 postdocs skilled in human cell culture. The Ray and Dagmar Dolby Regeneration Medicine Building, opened on 9 February, houses 25 labs and will spearhead UCSF stem-cell research. Recruitment will focus on blood-forming systems, transdifferentiation, pulmonary systems and skin, says Arnold Kriegstein, head of the UCSF stem-cell programme. The centre is one of 12 funded in part by the California Institute for Regenerative Medicine, a US\$3-billion state initiative approved by voters in 2004.

UNGLUED

A whisper in the dark.

BY AMBER D. SISTLA

So, it turns out that they don't make glue out of horse parts anymore. I take it off the list.

I still have another glue on the list — stopping parents who force their kids to sniff glue to keep them from getting drafted. Ms Tyler would like that one. She always says a mind is a terrible thing to waste. I also have gelatin, so I'd still be against misusing animal parts. Ms Tyler always says it's a shame to use animal parts for something frivolous like gelatin when so many people are starving, and starving wastes their minds.

With the other glue and the gelatin, the horse-glue was a redundant cause anyway. Dad always says reduce, reuse, remove redundancy. Said. Dad always said that. Before. I close my eyes tight and focus on my list. Crying wastes the mind, Ms Tyler says.

Glue and gelatin have always been on my list. I mean, glue is all sticky and gooey and it smells really good. And gelatin ... well, who doesn't like Jell-O? I mean, really.

I recite the list again and again in my head, waiting for the sounds to quieten on the other side of the privacy curtain. Whenever I start to get drowsy, I dig my nails into my legs.

Whisper won't come until I'm the only one still awake. Mum says Whisper is just a figment of my imagination, and I should ignore him. But if he's real, I think that would make him mad. Really mad. I'd rather be safe than sorry. That's what Dad always used to say.

Finally, all is quiet, and I peek through a small hole in the curtain. Mum's slumped against her workstation. Probably she'd been looking at her list — the one from the doctors with the cost of every procedure Dad needed. I'd heard Mum and Dad arguing because they only had enough money for one op.

I hear Whisper stirring, and I squeeze my eyes shut. Whisper always hides in the food processor on the bottom shelf. He slithers

and slides out of the appliance shelf, clattering across the shelf with the bowls and plates, scales scritch against the shelf with my robot, and then rubbing against the shelf with my clothes before finally arriving on top at my bed.

I don't open my eyes. Whisper would disappear. I know what I need to do, but the thought of doing it alone makes me feel like a used-up battery.

Are you ready?

He coils around me like a capacitor.

Whisper's question squirms in my mind like a live wire. I nod and open my mouth, and Whisper slides inside, tasting like lightning. He crams inside, pushing

Mum sleeps, curled around her station. Her black hair covers her face. She'll wake if I push it aside, so I just memorize its dark shadow.

I take one last look at our ten-by-ten. We used to have a ten-by-twenty when Mum and Dad still smiled. I don't remember that place much, but I do remember the smiles.

I slide the door open just enough to squeeze through. The halls are crowded. I asked Dad once why they didn't just get a one-by-one. He shrugged and said maybe they were doing a camping vacation, but Mum glared at him and said to Dad to be realistic and she told me that they didn't have money or the buildings just don't have room any more or both. When I looked up camping, it didn't make sense, so I guess Mum was right.

I wedge myself into the queue on the up stairway. On the 243rd floor, a lot of people exit, pushing and pulling me with them. My hands get clammy. This high up, it'd be impossible to get back on if I had a forced exit. Whisper growls, and suddenly there's enough room to keep squirming up and up to the top floor.

The recruitment centre is huge. Hundred-by-hundred, maybe. Bigger, maybe. It made me dizzy just to look at it, so I run around it as fast as I can until my legs won't move anymore.

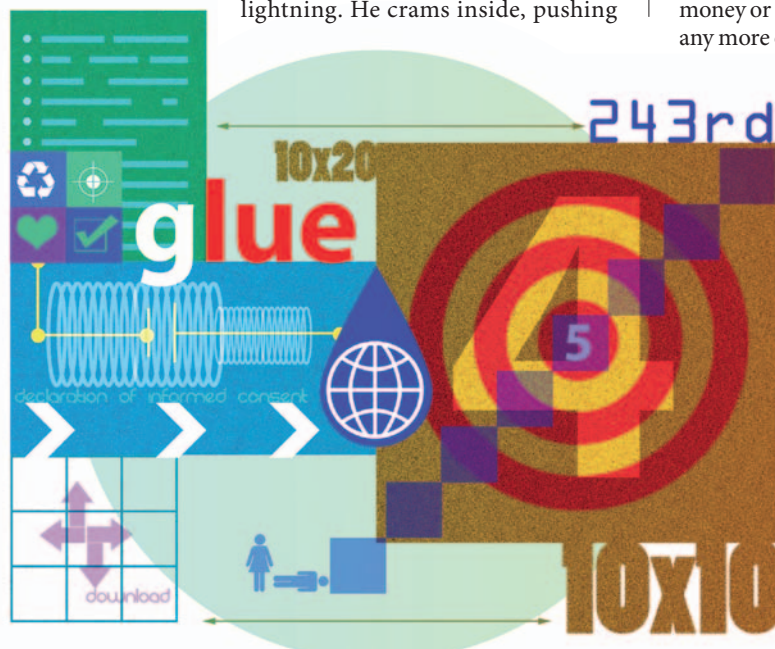
While I'm gasping for breath, the woman recites the Declaration of Informed Consent as if I'm a baby in a crèche. I've been hearing it every month at school for five years now, ever since I turned four and reached the age of Consent.

The woman tells me that she won't know yet what my brain will be used to intelligence until after the auction ends. After Consent Given, the base credit will transferred to my parents' account, and they could get a bonus depending on how the bidding goes.

She tells me that I'm helping save the world. I don't tell her about my list, or that I have my own plans.

Whisper says he has plans too. ■

Amber D. Sistla's fiction has appeared in Jim Baen's Universe and Cosmos magazine.



memories and thoughts about until I think my head will explode, but then it's done, and he's curled inside my mind.

I climb down, pausing to touch my robot. It's all ready, just needs a brain. Dad said if I built it and proved it worked to spec, he would get a mouse brain for me to intelligence it. But that was before.

Crying will ruin everything, so I think about my list until my tears forget they were coming. Down on the ground, I kneel by Dad's makeshift bed. He's deflated a lot since he got sick. I hold his hand and mouth the list in his ear. He doesn't move, the meds make sure of that. Nose to nose, I memorize his face.

► NATURE.COM
Follow Futures on
Facebook at:
go.nature.com/mtoodm

Coupled quantized mechanical oscillators

K. R. Brown¹, C. Ospelkaus¹, Y. Colombe¹, A. C. Wilson¹, D. Leibfried¹ & D. J. Wineland¹

The harmonic oscillator is one of the simplest physical systems but also one of the most fundamental. It is ubiquitous in nature, often serving as an approximation for a more complicated system or as a building block in larger models. Realizations of harmonic oscillators in the quantum regime include electromagnetic fields in a cavity^{1–3} and the mechanical modes of a trapped atom⁴ or macroscopic solid⁵. Quantized interaction between two motional modes of an individual trapped ion has been achieved by coupling through optical fields⁶, and entangled motion of two ions in separate locations has been accomplished indirectly through their internal states⁷. However, direct controllable coupling between quantized mechanical oscillators held in separate locations has not been realized previously. Here we implement such coupling through the mutual Coulomb interaction of two ions held in trapping potentials separated by 40 μm (similar work is reported in a related paper⁸). By tuning the confining wells into resonance, energy is exchanged between the ions at the quantum level, establishing that direct coherent motional coupling is possible for separately trapped ions. The system demonstrates a building block for quantum information processing and quantum simulation. More broadly, this work is a natural precursor to experiments in hybrid quantum systems, such as coupling a trapped ion to a quantized macroscopic mechanical or electrical oscillator^{9–13}.

The direct coupling of atomic ions in separate potential wells is a key feature of proposals to implement quantum simulation^{14–16}, and it could allow logic operations to be performed in a multi-zone quantum information processor^{10,17,18} without the requirement of bringing the ion qubits into the same trapping potential. Moreover, the coupling could prove useful for metrology and sensing. For example, it could extend the capabilities of quantum logic spectroscopy^{9,19,20} to ions that cannot be trapped within the same potential well as the measurement ion, such as oppositely charged ions or even antimatter particles^{9,10}. Coupling could be obtained either through mutually shared electrodes^{9,21} or directly through the Coulomb interaction^{10,17,22,23}.

The Coulomb interaction potential for two trapped ions, a and b , with charges q_a and q_b in potential wells separated by a distance s_0 is given by

$$U(x_a, x_b) = \frac{1}{4\pi\epsilon_0 s_0} \frac{q_a q_b}{-x_a + x_b} \\ \approx \frac{1}{4\pi\epsilon_0 s_0} \left(1 + \frac{x_a - x_b}{s_0} + \frac{x_a^2}{s_0^2} + \frac{x_b^2}{s_0^2} - \frac{2x_a x_b}{s_0^2} \right)$$

Here x_a and x_b are the displacements of the ions from the external potential minima and ϵ_0 is the permittivity of free space. The first term is constant and does not affect the dynamics. The second term represents a steady force between the ions that displaces them slightly; if necessary, it can be counteracted with additional potentials applied to nearby electrodes. The terms proportional to x_a^2 and x_b^2 represent static changes in the trap frequencies that could also be compensated with potentials applied to nearby electrodes. The term proportional to $x_a x_b$ represents the lowest-order coupling between the ions' motions. For small deviations, x'_a and x'_b , from equilibrium, the coupling is

$$\frac{-q_a q_b}{2\pi\epsilon_0 s_0^3} (x'_a x'_b) = -\hbar\Omega_{\text{ex}}(a + a^\dagger)(b + b^\dagger) \approx -\hbar\Omega_{\text{ex}}(ab^\dagger + a^\dagger b) \quad (1)$$

where

$$\Omega_{\text{ex}} \equiv \frac{q_a q_b}{4\pi\epsilon_0 s_0^3 \sqrt{m_a m_b} \sqrt{\omega_{0a} \omega_{0b}}} \quad (2)$$

and a, a^\dagger, b and b^\dagger represent the harmonic oscillator lowering (a, b) and raising (a^\dagger, b^\dagger) operators, m_i and ω_{0i} are respectively the ion masses and motional frequencies, \hbar is Planck's constant divided by 2π , and we have neglected fast-rotating terms. Minimizing the distance, s_0 , between ions is crucial, because for fixed values of ω_{0i} , the coupling rate scales as $\Omega_{\text{ex}} \propto 1/s_0^3$.

When $\omega_{0a} = \omega_{0b} = \omega_0$ (the resonance condition), we find that

$$a^\dagger(t) = \exp(i\omega_0 t) (a^\dagger(0) \cos(\Omega_{\text{ex}} t) - i b^\dagger(0) \sin(\Omega_{\text{ex}} t)) \\ b^\dagger(t) = \exp(i\omega_0 t) (b^\dagger(0) \cos(\Omega_{\text{ex}} t) - i a^\dagger(0) \sin(\Omega_{\text{ex}} t)) \quad (3)$$

At time $t = \tau_{\text{ex}} \equiv \pi/2\Omega_{\text{ex}}$, the operators have changed roles up to a phase factor, and the oscillators have completely swapped their energies, regardless of their initial states. At $t = 2\tau_{\text{ex}}$, the energies have returned to their initial values in each ion. The mean occupation, $\langle a^\dagger a \rangle$, of ion a as a function of time exhibits oscillations with period $2\tau_{\text{ex}}$.

Figure 1 shows a micrograph of our surface-electrode trap²⁴, constructed of gold electrodes, 8 μm thick with 5- μm gaps, electroplated

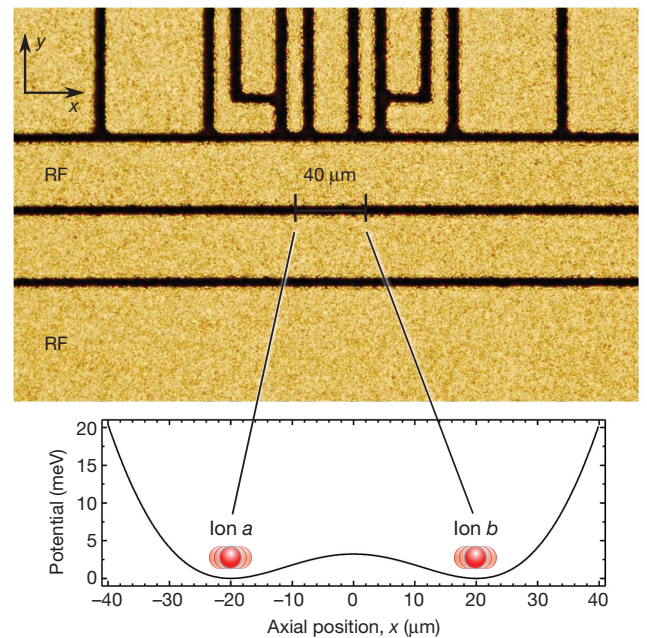


Figure 1 | Micrograph of the ion trap, showing radio-frequency (RF) and d.c. electrodes, and gaps between electrodes (darker areas). The lower part of the figure indicates the simulated potential along the trap x axis. Two trapping wells are separated by 40 μm , with ion positions marked by red spheres. The d.c. electrodes are sufficient to control the axial frequency and the position of each ion independently. Here both frequencies are ~ 4 MHz and the potential barrier between the two ions is ~ 3 meV.

¹Time and Frequency Division, National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA.

onto a crystalline quartz substrate. The trap can produce two potential minima at a height $d_0 = 40 \mu\text{m}$ above the surface and separated by $s_0 = 40 \mu\text{m}$ along the x axis. Each potential well confines a single $^9\text{Be}^+$ ion with an axial (parallel to x) motional frequency of $\omega_0/2\pi \approx 4 \text{ MHz}$ and a barrier between wells of $\sim 3 \text{ meV}$. Pseudopotential confinement in radial directions (normal to x) is accomplished with a peak potential of $\sim 100 \text{ V}$ at 170 MHz applied to the radio-frequency electrodes, yielding radial frequencies of $\sim 22 \text{ MHz}$. By applying static potentials to the d.c. electrodes, we can independently vary the separation between the ions and the curvatures of the two trapping wells. In this way, the ion axial motional frequencies can be brought into or out of resonance, allowing a tunable interaction. For $\omega_0/2\pi = 4.04 \text{ MHz}$, we predict that $\tau_{\text{ex}} = 162 \mu\text{s}$, where we have included a 2% correction in Ω_{ex} owing to the metallic electrodes beneath the ions (Methods Summary).

With currently achieved size scales in ion traps, the Coulomb interaction is relatively weak, so low ion heating rates and stable trapping potentials are essential. Heating rates can be suppressed by operating at cryogenic temperatures^{25,26}, such that the direct Coulomb coupling rate can exceed the heating rate. In this work, the trap electrodes and surrounding vacuum enclosure are cooled to 4.2 K with a liquid-helium bath cryostat. With similar versions of this trap at $\omega_0/2\pi = 2.3 \text{ MHz}$, heating rates expressed as $d\langle n \rangle/dt$ (where n denotes the quantum number of motional Fock state $|n\rangle$ and $\langle n \rangle$ is its expectation value) were observed to be as low as 70 quanta per second, consistent with the results of ref. 26. However, for the experiments described here ($\omega_0/2\pi \approx 4.0\text{--}5.6 \text{ MHz}$) the heating rate was $\sim 500\text{--}2,000$ quanta per second and varied between the two wells. We observed $d\langle n \rangle/dt \propto 1/\omega_0^2$ in this trap, in agreement with previous reports^{25–27}, so large values of ω_0 are beneficial. The use of $^9\text{Be}^+$, the lightest of the commonly trapped atomic ions, is an advantage here, because for given d.c. trapping potentials the heating rate should remain unchanged while $\Omega_{\text{ex}} \propto m^{-1/2}$. Cryogenic operation decreases the background gas pressure to negligible levels, such that ion loss rates due to collisions with background gas are smaller than one per day.

A signature of coupling between the ions is the splitting between the two axial normal mode frequencies. As the trap potential is tuned into the resonance condition, this splitting, δf , reaches a theoretical minimum of $\delta f = \Omega_{\text{ex}}/\pi = 3.1 \text{ kHz}$. A plot of the mode frequencies will therefore show an avoided crossing. We measure the mode frequencies by applying a nearly resonant oscillating potential pulse to one of the trap electrodes. We then illuminate both ions with laser radiation resonant with the $^2S_{1/2} \rightarrow ^2P_{3/2}$ cycling transition at 313 nm . A decrease in the resulting fluorescence indicates that a mode of the ions' motion has been resonantly excited. For pulse lengths $\tau_p \gg 1/\delta f$, we resolve the two modes (Fig. 2a). We sweep the trapping wells through resonance by varying the static potentials that are applied to the trap electrodes. A plot of the resulting mode frequencies, determined as above, is given in Fig. 2b, c, showing a minimum of $\delta f = 3.0(5) \text{ kHz}$, in agreement with theory.

To demonstrate coupling at the level of a few motional quanta, we observe the exchange of energy between the two ions as follows. The ion motional frequencies are initially detuned by 100 kHz , which is much greater than $\Omega_{\text{ex}}/2\pi$, effectively decoupling the ions' motions. The ions are then simultaneously illuminated with a laser detuned by -10 MHz from the $^2S_{1/2} \rightarrow ^2P_{3/2}$ cycling transition, cooling them into a thermal state at the Doppler limit with mean occupation $\langle n \rangle = 2.3(1)$. Subsequently, ion a is cooled to $\langle n_a \rangle = 0.35(2)$ by several cycles of stimulated Raman cooling with the $|\downarrow\rangle \equiv |F=2, m_F=-2\rangle$ and $|\uparrow\rangle \equiv |1, -1\rangle$ hyperfine states²⁸. The Raman beams are counter-propagating and oriented at 45° relative to the x axis. At this point, the potentials are brought into resonance ($\omega_0/2\pi = 4.04 \text{ MHz}$) within an interval ($9 \mu\text{s}$) short in comparison with τ_{ex} but long in comparison with the axial oscillation period. They remain on resonance for a time τ , allowing energy to transfer between the ions. After a time τ , the potentials are adiabatically returned to their off-resonance values and we determine the mean quantum number, $\langle n_a \rangle$, in ion a by observing asymmetry between the red and blue motional sidebands of the $|\downarrow\rangle$ -to- $|\uparrow\rangle$ hyperfine Raman transition²⁸.

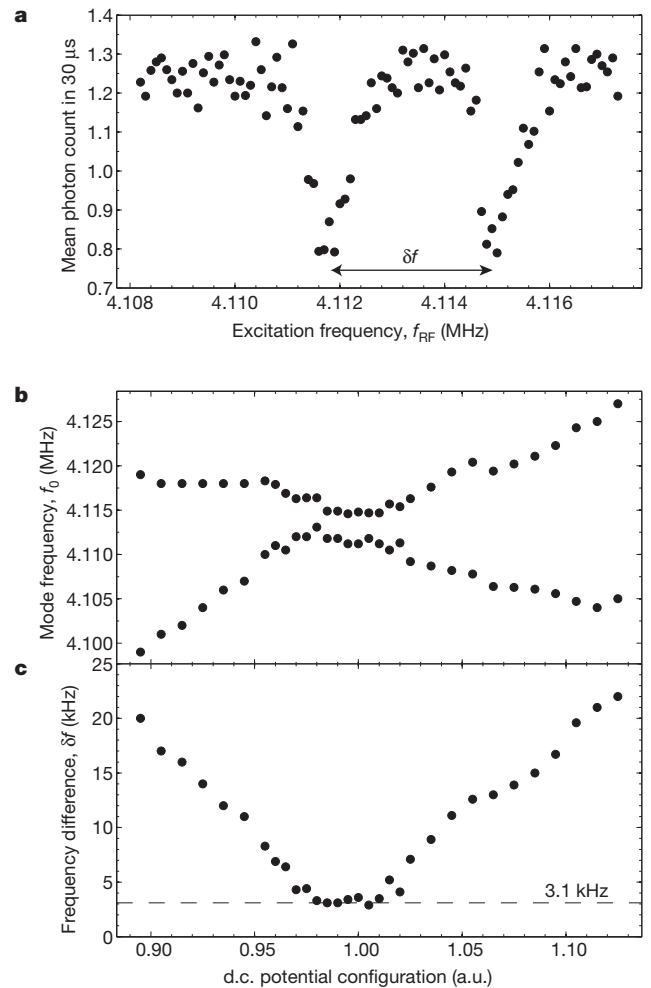


Figure 2 | Motional spectroscopy of two coupled ions near the avoided crossing. **a**, Decreases in collected fluorescence occur at values of excitation frequency, f_{RF} , corresponding to the ion mode frequencies. With $\tau_p = 960 \mu\text{s}$, the splitting on resonance is resolved. **b**, **c**, Mode frequencies (**b**) and mode frequency splitting, δf (**c**), for the axial normal modes of two ions separated by $40 \mu\text{m}$. Error bars are smaller than the size of the points. The data were acquired over a 1-h period, and slow variations in ambient potentials gave rise to the fluctuations.

As seen in Fig. 3, energy exchanges between the ions during an interval $\tau_{\text{ex}} = 155(1) \mu\text{s}$. The 5% disagreement between the measured and predicted ($162 \mu\text{s}$) values for τ_{ex} is probably due to uncertainty in the ion separation, s_0 (even a $1\text{-}\mu\text{m}$ uncertainty would account for the disagreement). The first maximum of $\langle n_a \rangle$ corresponds to the cooling limit of ion b ($\langle n_b(\tau=0) \rangle = 2.3(1)$ quanta). The underlying linear growth in $\langle n_a \rangle$ corresponds to heating of the ions at a rate of $\dot{n} = 1,885(10)$ quanta per second (Methods Summary).

As a final experiment, we demonstrate energy exchange at approximately the single-quantum level. Ideally, the experiment takes the following form. The ions are tuned to the resonance condition throughout and are initially Doppler-cooled. Ion a is Raman-cooled, sympathetically cooling ion b and thereby preparing the state $|0\rangle_a |\downarrow\rangle_a |0\rangle_b$. To create a single motional quantum, we drive ion a with a blue-sideband Raman π pulse (of duration $10 \mu\text{s}$, which is much less than τ_{ex}), creating the state $|1\rangle_a |\uparrow\rangle_a |0\rangle_b$. The system oscillates between $|1\rangle_a |\uparrow\rangle_a |0\rangle_b$ and $|0\rangle_a |\uparrow\rangle_a |1\rangle_b$ with period $2\tau_{\text{ex}}$. After a time τ , we drive ion a with another blue-sideband π pulse, conditionally flipping the spin from $|1\rangle_a |\uparrow\rangle_a$ to $|0\rangle_a |\downarrow\rangle_a$, dependent on the presence of a motional quantum in ion a . The final internal state probability will be given by $P(|\uparrow\rangle_a)(\tau) = \sin^2(\Omega_{\text{ex}}\tau)$. In practice, contrast in the oscillations (Fig. 4) is significantly reduced by

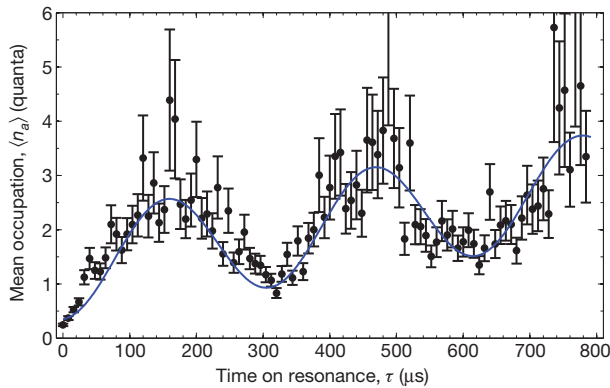


Figure 3 | Energy swapping between two ions in separate trapping potentials at the level of a few quanta. The mean occupation, $\langle n_a \rangle$, of ion a is plotted with error bars (s.e.m.) for various durations, τ , that the ion motional frequencies remain on resonance. The blue curve represents a fit to theory with four free parameters: the two initial mean quantum numbers, the exchange time and the heating rate. Energy exchanges between the ions at 155(1)- μs intervals. The linearly increasing trend in $\langle n_a \rangle$ is due to ion heating at a rate of 1,885(10) quanta per second. Uncertainties represent standard errors of the fit parameters.

incomplete cooling, motional decoherence and decoherence due to imperfect Raman sideband pulses. We estimate that ion a is cooled initially to $\langle n_a \rangle = 0.3(1)$. Although we were unable to measure the initial temperature of ion b directly, comparison of the contrast and temporal behaviour of our exchange data (Fig. 4) with simulations indicates that $\langle n_b \rangle \lesssim 0.6$. Motional decoherence results from heating and from trap frequency instability over the time required to acquire the data. Raman sideband pulses suffer from variations in laser intensity and fluctuations in the sideband coupling caused by thermal spread in the y and z motional states (Debye–Waller factors¹⁰). For $\omega_0/2\pi = 5.56$ MHz, we observe oscillations with period $2\tau_{\text{ex}} = 437(4)$ μs . The 2% disagreement with the prediction (447 μs) of equations (2) and (3) is probably due to uncertainty in the ion separation and the difficulty of maintaining the exact resonance condition.

Significant improvements in coupling fidelity seem to be in reach with current technology. It should be possible to improve Raman laser intensity stability, and Debye–Waller factors from the y and z motion can be eliminated by proper choice of beam directions²⁸. We believe that our motional frequency instability can be alleviated by ensuring that the trap surface is free of charged contaminants. Faster exchange can be achieved by scaling down trap dimensions, but this puts a premium on reduced motional heating. Although the work presented

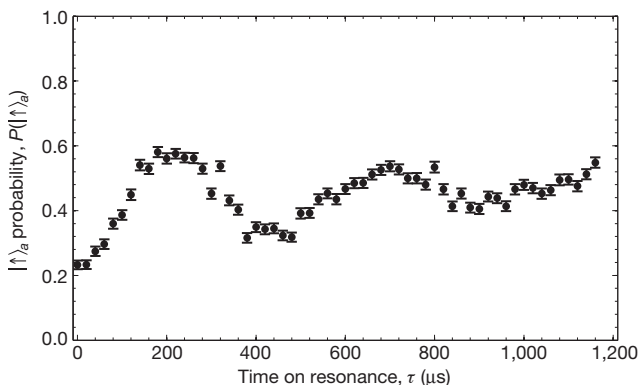


Figure 4 | Motional exchange between two ions in separate trapping potentials at approximately the single-quantum level. The probability, $P(|\uparrow\rangle_a)$, of measuring ion a in spin state $|\uparrow\rangle_a$ at the end of the experimental sequence is plotted with error bars (s.e.m.) against the time, τ , for which the ions interact. $P(|\uparrow\rangle_a)$ oscillates with period $2\tau_{\text{ex}} = 437(4)$ μs as a quantum exchanges between the ions.

here uses the axial mode to couple the ions, it may prove advantageous to use the radial modes in future experiments because of the lower heating rates associated with their higher frequencies. As an example, radial mode frequencies of ~ 30 MHz are routinely achieved in the apparatus, compared with axial frequencies of $\lesssim 10$ MHz.

These results could lead to several possible applications in quantum state engineering and spectroscopy. For example, from the motional state $|1\rangle_a|0\rangle_b$ the state of the ions at time $t = \tau_{\text{ex}}/2$ is the Bell state $(|1\rangle_a|0\rangle_b + i|0\rangle_a|1\rangle_b)/\sqrt{2}$. Transferring this state onto the ions' internal states with sideband pulses would create an entangled spin state, even between ions of dissimilar species. This could be used as an entangled-pair factory in the scheme of refs 10,18, with the advantage over previous schemes⁷ that the ions are already in separate wells, ready for distribution to separate locations. The coupling could be used to read out the state of one ion species with another, an ability useful for error correction protocols and for quantum logic spectroscopy¹⁹. When the harmonic wells are not in resonance, the spin state of one ion can be read out without destroying the state of the other, so schemes for weak or quantum non-demolition measurements that use either Kerr-type nonlinearities or quantum logic may become feasible¹. Hybrid quantum systems, incorporating similar interactions to couple a trapped ion to other quantum devices, could serve as a means of transferring quantum information between different qubit implementations in a future quantum network. For example, a trapped ion could act as a quantum transformer between a superconducting qubit¹³ and a photonic qubit^{29,30}. The sympathetic cooling through exchange might also be used to cool neutral molecules³¹.

METHODS SUMMARY

The shielding factor, β , represents the ratio of the exchange rates Ω_{ex} with and without the presence of the trapping electrodes. To a good approximation, we ensure that at the motional frequencies all trap electrodes are held at ground. Therefore, assuming that gaps between the electrodes are negligible, the shielding factor can be calculated with the method of images. The result is

$$\beta = 1 - \frac{1}{2} \left(\frac{3(s_0/d_0)^5}{(4 + (s_0/d_0)^2)^{5/2}} - \frac{(s_0/d_0)^3}{(4 + (s_0/d_0)^2)^{3/2}} \right)$$

which reaches a maximum of $\beta = 1.018$ at $s_0 = d_0$.

The evolution of $\langle n_a \rangle$ under the influence of equation (1), including heating effects and assuming that both ions begin in a thermal state with mean quantum numbers n_{a0} and n_{b0} , can be predicted with a Langevin equation⁹. On resonance, the evolution is

$$\langle n_a \rangle = n_{a0} \cos^2(\Omega_{\text{ex}}t) + n_{b0} \sin^2(\Omega_{\text{ex}}t) + \dot{n}t$$

where \dot{n} represents the mean of $d\langle n_a \rangle/dt$ and $d\langle n_b \rangle/dt$ for uncorrelated noise sources.

Received 9 November; accepted 1 December 2010.

Published online 23 February 2011.

1. Haroche, S. & Raimond, J.-M. *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford Univ. Press, 2006).
2. Miller, R. *et al.* Trapped atoms in cavity QED: coupling quantized light and matter. *J. Phys. B* **38**, S551 (2005).
3. Houck, A. A. *et al.* Generating single microwave photons in a circuit. *Nature* **449**, 328–331 (2007).
4. Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. J. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**, 281–324 (2003).
5. O'Connell, A. D. *et al.* Quantum ground state and single-phonon control of a mechanical resonator. *Nature* **464**, 697–703 (2010).
6. Monroe, C. *et al.* in *Atomic Physics 17* (eds Arimondo, E., De Natale, P. & Inguscio, M.) 173–186 (Proc. 17th Int. Conf. Atomic Phys., Springer, 2001).
7. Jost, J. D. *et al.* Entangled mechanical oscillators. *Nature* **459**, 683–685 (2009).
8. Harlander, M., Lechner, R., Brownnutt, M., Blatt, R. & Hänsel, W. Trapped-ion antennae for the transmission of quantum information. *Nature* doi:10.1038/nature09800 (this issue).
9. Heinzen, D. J. & Wineland, D. J. Quantum-limited cooling and detection of radio-frequency oscillations by laser-cooled ions. *Phys. Rev. A* **42**, 2977–2994 (1990).
10. Wineland, D. J. *et al.* Experimental issues in coherent quantum-state manipulation of trapped atomic ions. *J. Res. Natl. Inst. Stand. Technol.* **103**, 259–328 (1998).
11. Tian, L. & Zoller, P. Coupled ion-nanomechanical systems. *Phys. Rev. Lett.* **93**, 266403 (2004).

12. Hensinger, W. K. *et al.* Ion trap transducers for quantum electromechanical oscillators. *Phys. Rev. A* **72**, 041405(R) (2005).
13. Tian, L., Blatt, R. & Zoller, P. Scalable ion trap quantum computing without moving ions. *Eur. Phys. J. D* **32**, 201–208 (2005).
14. Schmied, R., Roscilde, T., Murg, V., Porras, D. & Cirac, J. I. Quantum phases of trapped ions in an optical lattice. *N. J. Phys.* **10**, 045017 (2008).
15. Chiaverini, J. & Lybarger, W. E. Jr. Laserless trapped-ion quantum simulations without spontaneous scattering using microtrap arrays. *Phys. Rev. A* **77**, 022324 (2008).
16. Schmied, R., Wesenberg, J. H. & Leibfried, D. Optimal surface-electrode trap lattices for quantum simulation with trapped ions. *Phys. Rev. Lett.* **102**, 233002 (2009).
17. Cirac, J. I. & Zoller, P. A scalable quantum computer with ions in an array of microtraps. *Nature* **404**, 579–581 (2000).
18. Kielpinski, D., Monroe, C. & Wineland, D. J. Architecture for a large-scale ion-trap quantum computer. *Nature* **417**, 709–711 (2002).
19. Schmidt, P. O. *et al.* Spectroscopy using quantum logic. *Science* **309**, 749–752 (2005).
20. Rosenband, T. *et al.* Frequency ratio of Al^+ and Hg^+ single-ion optical clocks; metrology at the 17th decimal place. *Science* **319**, 1808–1812 (2008).
21. Daniilidis, N., Lee, T., Clark, R., Narayanan, S. & Häffner, H. Wiring up trapped ions to study aspects of quantum information. *J. Phys. B* **42**, 154012 (2009).
22. Tan, J. N. Interacting ion oscillators in contiguous confinement wells. *Bull. Am. Phys. Soc.* **47**, 103 (2002).
23. Ciaramicoli, G., Marzoli, I. & Tombesi, P. Scalable quantum processor with trapped electrons. *Phys. Rev. Lett.* **91**, 017901 (2003).
24. Seidelin, S. *et al.* Microfabricated surface-electrode ion trap for scalable quantum information processing. *Phys. Rev. Lett.* **96**, 253003 (2006).
25. Deslauriers, L. *et al.* Scaling and suppression of anomalous heating in ion traps. *Phys. Rev. Lett.* **97**, 103007 (2006).
26. Labaziewicz, J. *et al.* Temperature dependence of electric field noise above gold surfaces. *Phys. Rev. Lett.* **101**, 180602 (2008).
27. Epstein, R. J. *et al.* Simplified motional heating rate measurements of trapped ions. *Phys. Rev. A* **76**, 033411 (2007).
28. Monroe, C. *et al.* Resolved-sideband Raman cooling of a bound atom to the 3D zero-point energy. *Phys. Rev. Lett.* **75**, 4011–4014 (1995).
29. Cirac, J. I., Zoller, P., Kimble, H. J. & Mabuchi, H. Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.* **78**, 3221–3224 (1997).
30. Moehring, D. L. *et al.* Entanglement of single-atom quantum bits at a distance. *Nature* **449**, 68–71 (2007).
31. Idziaszek, Z., Calarco, T. & Zoller, P. Ion-assisted ground-state cooling of a trapped polar molecule. Preprint at (<http://arxiv.org/abs/1008.1858>) (2010).

Acknowledgements This work was supported by IARPA, DARPA, ONR and the NIST Quantum Information Program. We thank M. Biercuk, A. VanDevender, J. Amini, and R. B. Blakestad for their help in assembling parts of the experiment, and we thank U. Warring and R. Simmonds for comments. This paper, a submission of NIST, is not subject to US copyright.

Author Contributions K.R.B. and C.O. participated in the design of the experiment and built the experimental apparatus. K.R.B. collected data, analysed results and wrote the manuscript. Y.C. fabricated the ion trap chip and collected data. A.C.W. maintained laser systems and collected data. D.L. participated in the design of the experiment, collected data and maintained laser systems. D.J.W. participated in the design and analysis of the experiment. All authors discussed the results and the text of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.R.B. (kenton.brown@nist.gov).

Structural basis of RNA polymerase II backtracking, arrest and reactivation

Alan C. M. Cheung¹ & Patrick Cramer¹

During gene transcription, RNA polymerase (Pol) II moves forwards along DNA and synthesizes messenger RNA. However, at certain DNA sequences, Pol II moves backwards, and such backtracking can arrest transcription. Arrested Pol II is reactivated by transcription factor IIS (TFIIS), which induces RNA cleavage that is required for cell viability¹. Pol II arrest and reactivation are involved in transcription through nucleosomes^{2,3} and in promoter-proximal gene regulation^{4–6}. Here we present X-ray structures at 3.3 Å resolution of an arrested *Saccharomyces cerevisiae* Pol II complex with DNA and RNA, and of a reactivation intermediate that additionally contains TFIIS. In the arrested complex, eight nucleotides of backtracked RNA bind a conserved ‘backtrack site’ in the Pol II pore and funnel, trapping the active centre trigger loop and inhibiting mRNA elongation. In the reactivation intermediate, TFIIS locks the trigger loop away from backtracked RNA, displaces RNA from the backtrack site, and complements the polymerase active site with a basic and two acidic residues that may catalyse proton transfers during RNA cleavage. The active site is demarcated from the backtrack site by a ‘gating tyrosine’ residue that probably delimits backtracking. These results establish the structural basis of Pol II backtracking, arrest and reactivation, and provide a framework for analysing gene regulation during transcription elongation.

Backtracking of bacterial RNA polymerase and the related eukaryotic Pol II to an arrested state is triggered by a weak DNA–RNA hybrid, and dislodges the RNA 3′ end from the active site^{7–9}. A recent study attempted to crystallize an arrested Pol II complex with the use of DNA–RNA scaffolds containing 3′-overhanging RNA¹⁰. This allowed the visualization of one or two backtracked RNA nucleotides, but not further-backtracked RNA¹⁰. We have now resolved further-backtracked RNA nucleotides by crystallization of an arrested complex obtained by Pol II transcription of a 3′-tailed DNA template. Tailed template transcription was previously used for structural studies¹¹, and is prone to arrest¹². We incubated Pol II with a tailed template (Fig. 1a) and different NTP substrates, crystallized the obtained complexes, collected diffraction data, and inspected difference electron density maps after phasing with the free Pol II structure. Only incubation with CTP alone resulted in interpretable difference density for backtracked RNA (Fig. 1b). The register of the nucleic acids was defined by bromine labelling of the –5 template base (Fig. 1a, b) and the structure was refined at 3.3 Å resolution (Methods and Supplementary Table 1).

The structure revealed 13 base pairs (bp) of downstream DNA, a 6-bp hybrid, and 9 nucleotides of single-stranded 3′ RNA that is extruded through the pore into the funnel (Fig. 1). Observation of a 15-nucleotide RNA with 9 unpaired residues is consistent with arrest of tailed template transcription after synthesis of 13–17 nucleotides¹², with a binding site on Pol II for a 9-nucleotide 3′ RNA¹³, and with patterns of RNA cleavage and nuclease protection in arrested complexes^{14,15}. Thus, the template had allowed for CMP mis-incorporation, and the resulting destabilized hybrid drove backtracking to the arrested state.

Comparison with the elongation complex structure reveals that the DNA–RNA hybrid is tilted towards the bridge helix (Fig. 1c). In the

elongation complex, the RNA 3′ end occupies position –1, and the incoming NTP substrate and the templating DNA base occupy position +1, relative to the active site. In the arrested complex, the –1 base pair is tilted by ~25°. Its DNA base resides in the site that is normally occupied by the +1 nucleotide, which instead resides in the downstream cleft, leaving the non-complementary +1 RNA base unpaired. In contrast, the +1 RNA base is paired with the DNA template base in a previously reported complex backtracked by one position¹⁰. Whether hybrid tilting is a cause or consequence of backtracking or whether it results from the shorter hybrid remains to be investigated.

The structure shows that backtracked RNA is extruded into the pore and funnel (Fig. 1d), confirming an early hypothesis¹⁶. The key feature of the complex is the highly defined structure of backtracked RNA and its extended binding site (Fig. 2). Backtracked RNA binds a ‘backtrack site’ along one side of the pore and the mobile trigger loop on the opposite side. The trigger loop has been proposed to control the lateral oscillation of polymerase¹⁷. Its ‘trapped’ conformation observed here is distinct from the closed, open and wedged conformations that occur during forward elongation^{18–21} (Supplementary Fig. 1), and also different from the conformation observed in a complex backtracked by one nucleotide¹⁰. These observations indicate the basis for transcription arrest. When backtracking is not extensive, RNA interactions with the backtrack site are partial and weak, and Pol II can spontaneously move forward and continue elongation. When backtracking is more extensive, backtracked RNA and the trigger loop are trapped, preventing forward elongation and arresting Pol II.

The structure reveals details of backtracked RNA and its interactions with the backtrack site (Fig. 2 and Supplementary Table 2). The first backtracked RNA nucleotide +2 contacts the bridge helix residue T827 and the fork residue E529, and stacks between the +1 base and Rpb2 residue Y769 that we call the ‘gating tyrosine’. The +2 base also contacts the gating tyrosine in a previous structure, although via an edge-to-face interaction¹⁰. Beyond position +2, base stacking discontinues due to steric hindrance by the gating tyrosine. The RNA backbone kinks between nucleotides +2 and +3, and contacts Rpb2 residue R766. The +3 nucleotide binds the trigger loop residues Q1078 and T1080, consistent with crosslinking data²². Beyond position +3, RNA binds exclusively to the Rpb1 funnel domain. A stack of bases from nucleotides +3 to +5 is followed by nucleotide +6, which inserts its base into a pocket that is flanked by R726 and I756 and called here the ‘funnel pocket’. Between positions +6 and +7, the RNA kinks again. The +7 base stacks against residue F755 on one side and against the +8 and +9 bases on the other. Because RNA-binding residues are conserved among eukaryotes (Supplementary Table 2 and Supplementary Fig. 2), backtracked RNA probably binds the same backtrack site in all Pol II enzymes. Modelling purine bases onto the backtracked RNA reveals potential clashes with Pol II (Supplementary Fig. 3). This indicates that the backtrack site preferentially accommodates pyrimidine bases, and provides an explanation for why known DNA arrest sites direct synthesis of pyrimidine-rich RNA²³.

The RNA conformation and Pol II interactions indicate why intrinsic RNA cleavage generally occurs after backtracking by one position, a

¹Gene Center and Department of Biochemistry, Center for Integrated Protein Science Munich, Ludwig-Maximilians-Universität München, Feodor-Lynen-Str. 25, 81377 Munich, Germany.

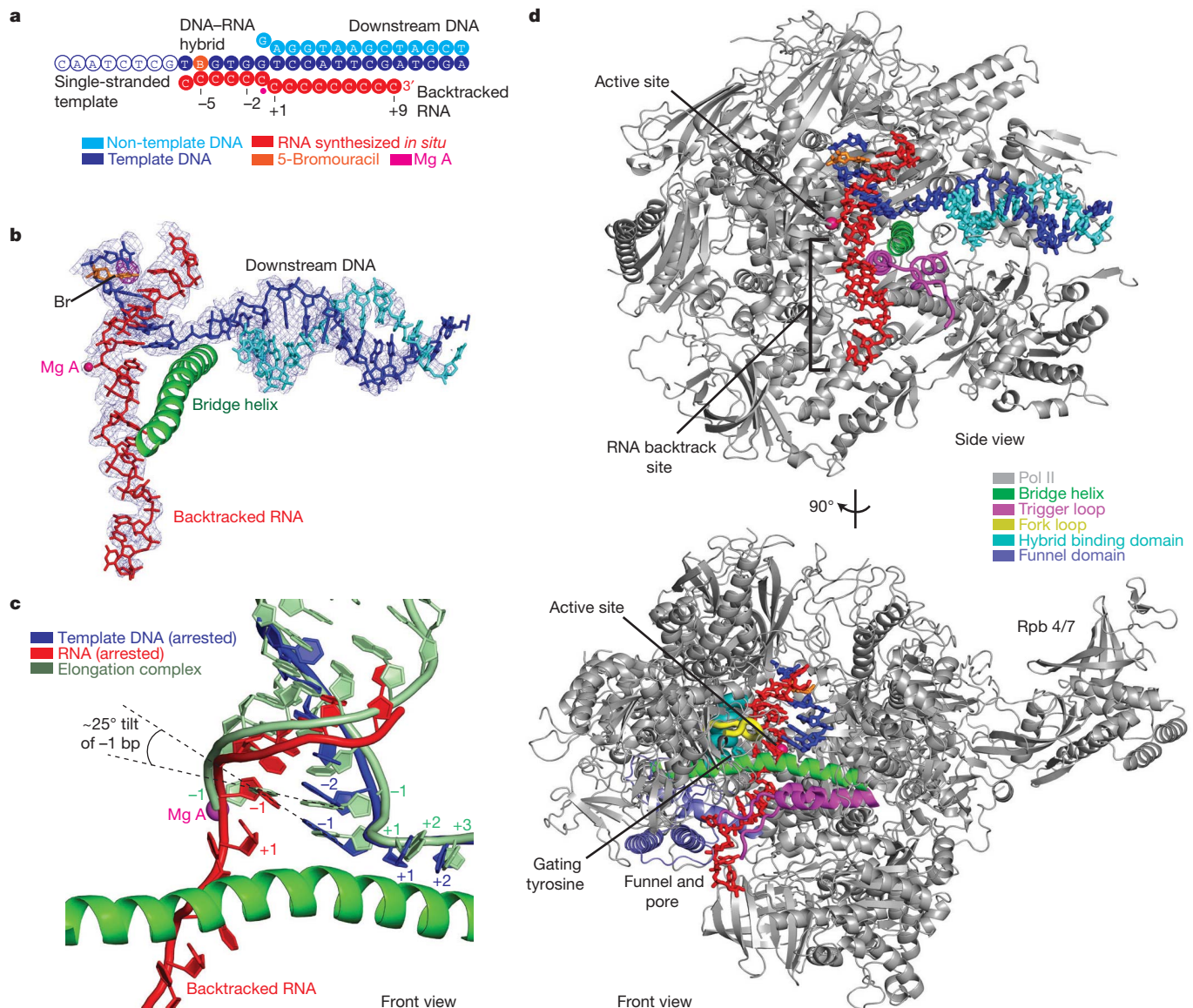


Figure 1 | Structure of arrested Pol II. **a**, Schematic of nucleic acids. Ordered nucleotides are shown with filled circles. The colour code is used throughout. The NTP-binding site corresponds to position +1 and backtracked (downstream) residues are labelled with positive numbers. Some complexes may contain additional disordered nucleotides at the RNA 3' end. **b**, Unbiased difference electron density (blue mesh, contoured at 3.0σ) for nucleic acids,

phenomenon that underlies mRNA proofreading^{10,24,25}. Backtracking begins with Pol II pausing and fraying of the 3'-terminal RNA nucleotide +1 against the gating tyrosine^{24,26}. The gating tyrosine maintains contact with the 3' nucleotide during the first step of backtracking¹⁰ (Fig. 2a). Backtracking by one position may thus be facilitated, but further backtracking is probably disfavoured because RNA base stacking must be discontinued at the gating tyrosine. Hence, the gating tyrosine may generally delimit the extent of backtracking. If, however, base-stacking interactions and the hybrid are weak, backtracking beyond the gating tyrosine may occur and lead to arrest.

Arrested Pol II is reactivated by TFIIS, which induces cleavage of backtracked RNA. TFIIS binds with its domain II near the rim of the Pol II funnel, and extends into the Pol II pore with its domain III, which reaches the active site with a β -stranded hairpin^{19,27}. Superposition of the arrested complex with the Pol II-TFIIS complex²⁷ revealed that the backtracked RNA overlaps with TFIIS domain III. This indicated that backtracked RNA prevents TFIIS from invading the pore, posing the

question of how reactivation occurs. To investigate this, we soaked arrested complex crystals with a TFIIS variant that carries two point mutations in functionally essential acidic hairpin residues (D290A/E291A), and solved the structure of the resulting reactivation intermediate at 3.3 Å resolution (Methods). As observed before¹⁹, TFIIS changed the Pol II conformation, realigned RNA in the hybrid, inserted its domain III into the pore, and reached the active site with its hairpin.

In the structure of the reactivation intermediate, TFIIS binding moved the trigger loop by up to 5 Å from the trapped to the 'locked' position (Fig. 3a) observed previously²⁷. TFIIS binding also induces a rotation of the gating tyrosine side chain, which prevents its stacking with backtracked RNA. Backtracked RNA was displaced from the backtrack site into the part of the pore that remains accessible after binding of TFIIS domain III ('restricted pore', Fig. 3b). The electron density for displaced RNA was discontinuous and only allowed for backbone modelling (Fig. 3b). The discontinuous density apparently

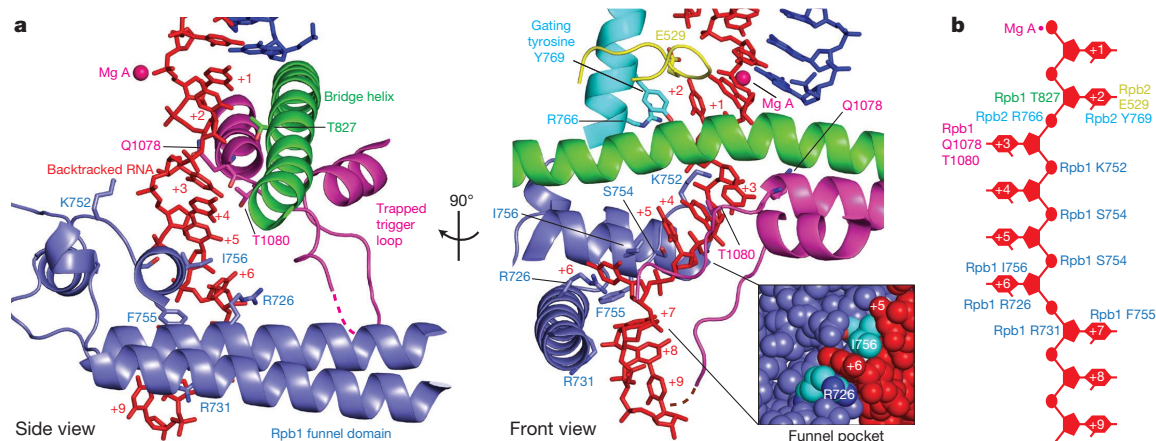


Figure 2 | Backtracked RNA in the backtrack site. **a**, Side and front views of backtracked RNA. RNA-binding Pol II elements in the pore and funnel are depicted, and contact residues in the backtrack site are labelled. **b**, Schematic of Pol II interactions with backtracked RNA.

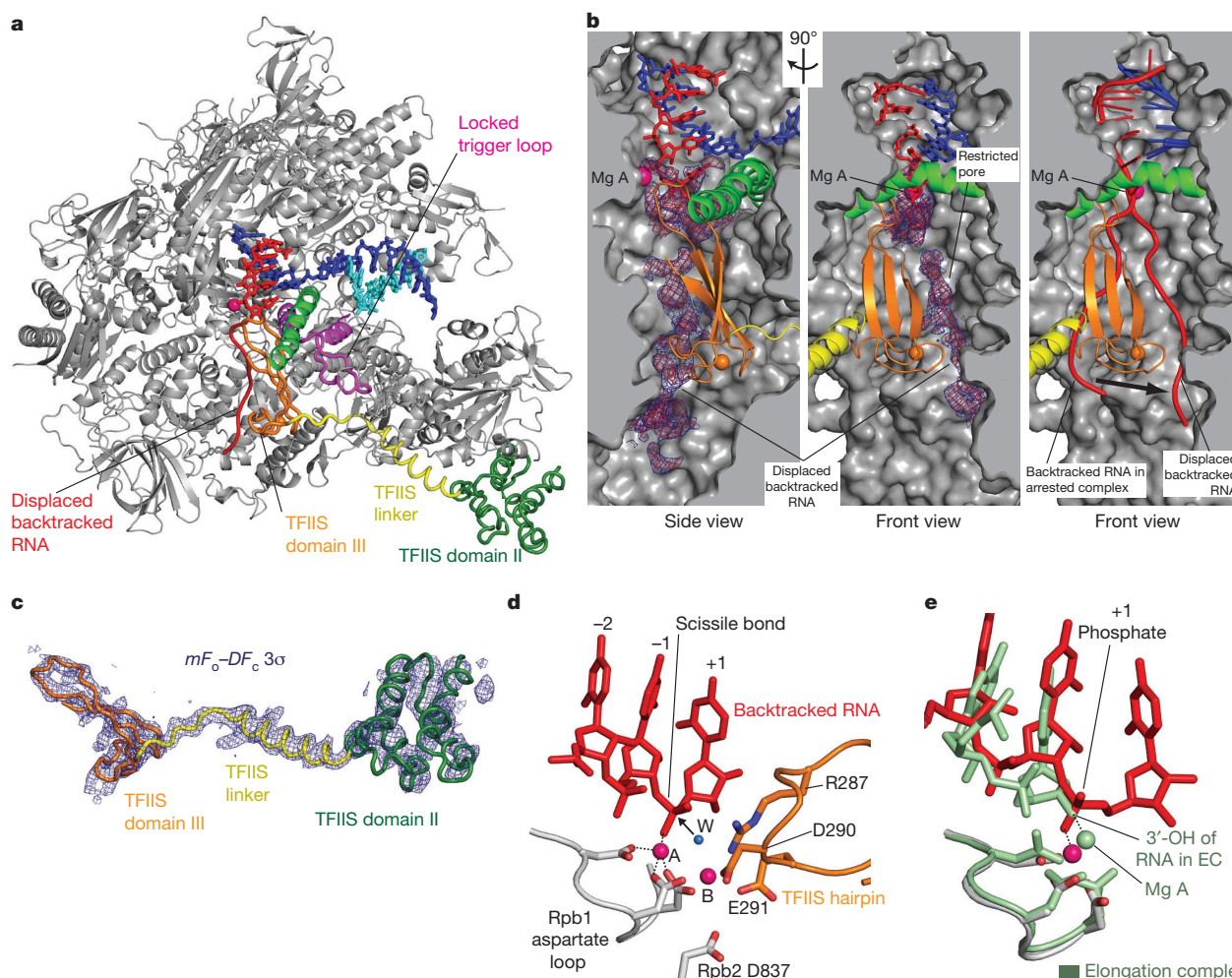


Figure 3 | Structure of reactivation intermediate. **a**, Side view. TFIIS domains II and III are in green and orange, respectively, and the domain II–III linker is in yellow. **b**, Displacement of backtracked RNA from the backtrack site into the restricted pore. Difference electron density for displaced backtracked RNA is contoured at 2σ (blue) and 2.5σ (red). On the right, backtracked RNA from the arrested complex is modelled into the structure, revealing a clash with TFIIS domain III. **c**, Unbiased difference electron density for TFIIS after phasing with the Pol II structure contoured at 3σ , and sigmaA-weighted by

coefficients m and D . **d**, Model for active site geometry during TFIIS-induced RNA cleavage. Metal A is coordinated by the Pol II aspartate loop and the RNA +1 phosphate. Metal B, the nucleophilic water molecule (W, blue sphere), and side chains of TFIIS hairpin residues D290 and E291 (orange) are placed onto the crystal structure without clashes. This supports a catalytically competent arrangement for an S_N2 mechanism. The arrow indicates the direction of the in-line nucleophilic attack. **e**, Superposition of the view in **d** with the active site conformation of the Pol II elongation complex¹⁹.

reflects rotational mobility rather than partial release of backtracked RNA, as RNA was not cleaved in a reconstituted backtracked complex by the inactive TFIIIS variant (not shown). These observations indicate that TFIIIS weakens the Pol II grip on backtracked RNA, and displaces and mobilizes RNA by competitive binding to the backtrack site.

Pol II reactivation by TFIIIS-dependent cleavage apparently involves two metal ions and a nucleophilic water molecule^{27–29}. Metal A is persistently bound to the active site aspartate loop¹⁶, whereas metal B may be recruited by TFIIIS^{27,29}. The reactivation intermediate structure provides new insights into TFIIIS-induced cleavage, as it is at higher resolution than previous work^{10,19}, and because it reveals the RNA +1 nucleotide, and thus the scissile phosphodiester bond between nucleotide –1 and +1 (Fig. 3d). First, metal A binds the +1 RNA phosphate to align the scissile bond, in contrast to its binding of the RNA 3' hydroxyl during nucleotide addition (Fig. 3e). Second, the TFIIIS hairpin residue R287 reaches the catalytic site and could stabilize the negatively charged transition state, explaining its role in catalysis³⁰. Third, modelling the side chains of D290 and E291 onto the structure

indicates that they can coordinate metal B, together with the Rpb2 residue D837 (Fig. 3d). The invariant charged TFIIIS residues R287, D290 and E291 may be required for two catalytic proton transfers, proton subtraction from the nucleophilic water, and proton donation to the product RNA 3' terminus.

The presented structural snapshots of transcription intermediates reveal the mechanisms of Pol II backtracking, arrest and reactivation (Fig. 4). When Pol II encounters a DNA sequence that impairs elongation, it pauses and backtracks by one position. Because further backtracking is hindered by the gating tyrosine, polymerase-intrinsic cleavage of a dinucleotide from the RNA 3' end can occur, and elongation continues. However, at an arrest site the hybrid is weak and the RNA can backtrack beyond the gating tyrosine. Extensive backtracking traps RNA and the trigger loop in the pore, inhibiting elongation and arresting Pol II. TFIIIS reactivates arrested Pol II by locking the trigger loop away from RNA, displacing and mobilizing backtracked RNA in the pore, and complementing the active site with a basic and two acidic side chains. This induces cleavage and release of backtracked RNA, and creates a new RNA 3' end at the active site that allows transcription to resume.

METHODS SUMMARY

Saccharomyces cerevisiae 12-subunit Pol II was prepared as described²⁴. Purified Pol II (3.5 mg ml^{–1}) was mixed with a twofold molar excess of tailed template (Fig. 1a) prepared as described¹⁹, 8 mM magnesium chloride and 2 mM CTP, and incubated for 2 h at 20 °C before crystallization by vapour diffusion with 6% PEG 6000, 200 mM ammonium acetate, 300 mM sodium acetate, 50 mM HEPES pH 7.0 and 5 mM TCEP as reservoir solution. Crystals were grown for 5–10 days, cryo-protected in mother solution supplemented with 22% glycerol and containing 2 mM CTP, 8 mM magnesium chloride, and 4 μM tailed template, followed by overnight incubation at 8 °C before harvesting and freezing in liquid nitrogen. Arrested Pol II–TFIIIS complex crystals were prepared by adding the inactive TFIIIS variant D290A/E291A²⁷ to the cryo-protectant at 1 mg ml^{–1} and incubating arrested Pol II crystals overnight at 8 °C. Diffraction data at 3.3 Å were collected at 100 K at beamline X06SA of the Swiss Light Source and structures were solved with molecular replacement using the model of 12-subunit Pol II (1WCM). Data were collected at 13.494 keV, the K-absorption peak of bromine.

Received 13 October; accepted 23 December 2010.

Published online 23 February 2011.

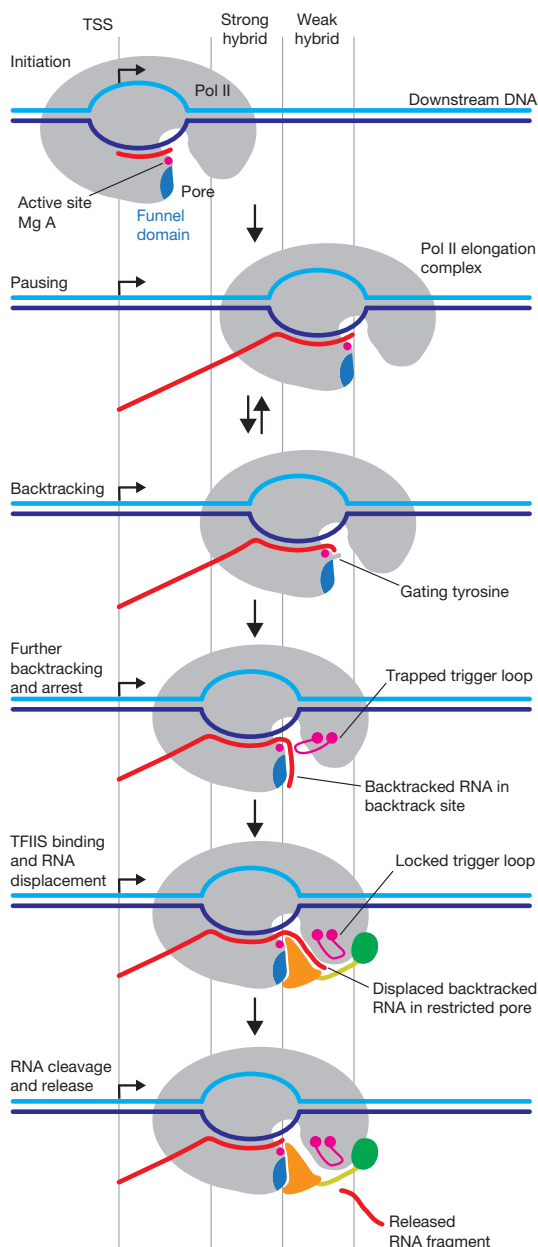


Figure 4 | Mechanism of Pol II backtracking, arrest and reactivation. Schematic of Pol II functional states. For details, compare text.

1. Sigurdsson, S., Dirac-Svejstrup, A. B. & Svejstrup, J. Q. Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. *Mol. Cell* **38**, 202–210 (2010).
2. Kim, J., Guermah, M. & Roeder, R. G. The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIIS. *Cell* **140**, 491–503 (2010).
3. Kireeva, M. L. *et al.* Nature of the nucleosomal barrier to RNA polymerase II. *Mol. Cell* **18**, 97–108 (2005).
4. Adelman, K. *et al.* Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIIS. *Mol. Cell* **17**, 103–112 (2005).
5. Nechaev, S. *et al.* Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* **327**, 335–338 (2010).
6. Palangat, M., Renner, D. B., Price, D. H. & Landick, R. A negative elongation factor for human RNA polymerase II inhibits the anti-arrest transcript-cleavage factor TFIIIS. *Proc. Natl Acad. Sci. USA* **102**, 15036–15041 (2005).
7. Komissarova, N. & Kashlev, M. RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. *J. Biol. Chem.* **272**, 15329–15338 (1997).
8. Nudler, E., Mustaev, A., Lukhtanov, E. & Goldfarb, A. The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* **89**, 33–41 (1997).
9. Palangat, M. & Landick, R. Roles of RNA:DNA hybrid stability, RNA structure, and active site conformation in pausing by human RNA polymerase II. *J. Mol. Biol.* **311**, 265–282 (2001).
10. Wang, D. *et al.* Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* **324**, 1203–1206 (2009).
11. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* **292**, 1876–1882 (2001).
12. Sluder, A. E., Price, D. H. & Greenleaf, A. L. Elongation by *Drosophila* RNA Polymerase-II. Transcription of 3'-extended DNA templates. *J. Biol. Chem.* **263**, 9917–9925 (1988).
13. Johnson, T. L. & Chamberlin, M. J. Complexes of yeast RNA polymerase II and RNA are substrates for TFIIIS-induced RNA cleavage. *Cell* **77**, 217–224 (1994).

14. Gu, W., Powell, W., Mote, J. J. & Reines, D. Nascent RNA cleavage by arrested RNA polymerase II does not require upstream translocation of the elongation complex on DNA. *J. Biol. Chem.* **268**, 25604–25616 (1993).
15. Izban, M. G. & Luse, D. S. The increment of SII-facilitated transcript cleavage varies dramatically between elongation competent and incompetent RNA polymerase II ternary complexes. *J. Biol. Chem.* **268**, 12874–12885 (1993).
16. Cramer, P. *et al.* Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**, 640–649 (2000).
17. Bar-Nahum, G. *et al.* A ratchet mechanism of transcription elongation and its control. *Cell* **120**, 183–193 (2005).
18. Brueckner, F. & Cramer, P. Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nature Struct. Mol. Biol.* **15**, 811–818 (2008).
19. Kettenberger, H., Armache, K.-J. & Cramer, P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIIS. *Mol. Cell* **16**, 955–965 (2004).
20. Vassylyev, D. G. *et al.* Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**, 163–168 (2007).
21. Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954 (2006).
22. Markovtsov, V., Mustaev, A. & Goldfarb, A. Protein-RNA interactions in the active center of transcription elongation complex. *Proc. Natl Acad. Sci. USA* **93**, 3221–3226 (1996).
23. Hawryluk, P. J., Ujvari, A. & Luse, D. S. Characterization of a novel RNA polymerase II arrest site which lacks a weak 3' RNA-DNA hybrid. *Nucleic Acids Res.* **32**, 1904–1916 (2004).
24. Sydow, J. F. *et al.* Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol. Cell* **34**, 710–721 (2009).
25. Zenkin, N., Yuzenkova, Y. & Severinov, K. Transcript-assisted transcriptional proofreading. *Science* **313**, 518–520 (2006).
26. Touloukhanov, I., Zhang, J. W., Palangat, M. & Landick, R. A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol. Cell* **27**, 406–419 (2007).
27. Kettenberger, H., Armache, K.-J. & Cramer, P. Architecture of the RNA polymerase II-TFIIIS complex and implications for mRNA cleavage. *Cell* **114**, 347–357 (2003).
28. Sosunov, V. *et al.* Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J.* **22**, 2234–2244 (2003).
29. Weillbaecher, R. G., Awrey, D. E., Edwards, A. M. & Kane, C. M. Intrinsic transcript cleavage in yeast RNA polymerase II elongation complexes. *J. Biol. Chem.* **278**, 24189–24199 (2003).
30. Awrey, D. E. *et al.* Yeast transcript elongation factor TFIIIS, structure and function. II: RNA polymerase binding, transcript cleavage, and read-through. *J. Biol. Chem.* **273**, 22595–22605 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank F. Brueckner, G. Damsma, K. Kinkelin, D. Kostrewa, L. Larivière, E. Lehmann, F. Martinez, S. Sainsbury and J. Sydow. Part of this work was performed at the Swiss Light Source at the Paul Scherrer Institut, Villigen, Switzerland. P.C. was supported by the Deutsche Forschungsgemeinschaft, SFB646, TR5, FOR1068, NIM, Bioimaging Network BIN, and the Jung-Stiftung.

Author Contributions A.C.M.C. carried out experiments. P.C. supervised the project. A.C.M.C. and P.C. prepared the manuscript.

Author Information Coordinates and structure factors of the arrested Pol II elongation complex and the arrested Pol II reactivation intermediate have been deposited with the Protein Data Bank under accession numbers 3PO2 and 3PO3, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.C. (cramer@lmb.uni-muenchen.de).

Trapped-ion antennae for the transmission of quantum information

M. Harlander¹, R. Lechner¹, M. Brownnutt¹, R. Blatt^{1,2} & W. Hänsel^{1,2†}

More than 100 years ago, Hertz succeeded in transmitting signals over a few metres to a receiving antenna using an electromagnetic oscillator, thus proving the electromagnetic theory¹ developed by Maxwell. Since this seminal work, technology has developed, and various oscillators are now available at the quantum mechanical level. For quantized electromagnetic oscillations, atoms in cavities can be used to couple electric fields^{2,3}. However, a quantum mechanical link between two mechanical oscillators (such as cantilevers^{4,5} or the vibrational modes of trapped atoms⁶ or ions^{7,8}) has been rarely demonstrated and has been achieved only indirectly. Examples include the mechanical transport of atoms carrying quantum information⁹ or the use of spontaneously emitted photons¹⁰. Here we achieve direct coupling between the motional dipoles of separately trapped ions over a distance of 54 micrometres, using the dipole–dipole interaction as a quantum mechanical transmission line¹¹. This interaction is small between single trapped ions, but the coupling is amplified by using additional trapped ions as antennae. With three ions in each well, the interaction is increased by a factor of seven compared to the single-ion case. This enhancement facilitates bridging of larger distances and relaxes the constraints on the miniaturization of trap electrodes. The system provides a building block for quantum computers and opportunities for coupling different types of quantum systems.

The exchange of quantum information between quantum bits (qubits) at remote sites is a key feature required to render quantum computation truly scalable¹². The dipole–dipole interaction offers a link between separate quantum systems without the need to shuttle particles between sites¹³. The interaction strength depends on the orientation and distance of the dipoles, and is in general given by

$$U_{dd} = \frac{1}{4\pi\epsilon_0} \frac{\mathbf{d}_1 \mathbf{d}_2 - 3(\mathbf{d}_1 \mathbf{e}_r)(\mathbf{d}_2 \mathbf{e}_r)}{r^3} \quad (1)$$

where \mathbf{d}_i are the interacting dipoles, r and \mathbf{e}_r denote respectively the magnitude and direction of their separation, and ϵ_0 is the vacuum permittivity.

Here this interaction is explored using ions or ion strings held in two separate potential wells of a linear segmented ion trap (Fig. 1), where the interacting dipoles are produced by the oscillating charges. As the dipole–dipole interaction decreases rapidly with trap separation, it is advantageous to bring the trapping wells as close together as possible. However, the generation of small inter-well distances requires similarly small distances, d , between the ions and the trap electrodes. This requirement runs counter to the effort to keep a larger ion–electrode separation in order to reduce both ‘anomalous heating’^{14,15} with its d^{-4} scaling and the effects of technical noise of the applied voltages¹⁶. Various routes may be taken to balance these competing requirements. In one approach, the heating rate may be reduced through the use of cryogenic temperatures¹⁷. Here, another approach is taken that uses more ions in the individual traps and enables interaction over larger distances. The additional ions work as ‘antennae’ that increase the motional dipole moment at the respective trapping site.

Given longitudinal alignment of the traps with one particle in each well, the dipole–dipole interaction (equation (1)) is:

$$U_{dd} = -\frac{q_1 q_2}{2\pi\epsilon_0} \frac{\Delta z_1 \Delta z_2}{r^3} \quad (2)$$

$$= -\hbar \frac{\Omega_c}{2} (a_1 + a_1^\dagger)(a_2 + a_2^\dagger) \\ \approx -\hbar \frac{\Omega_c}{2} (a_1 a_2^\dagger + a_1^\dagger a_2) \quad (3)$$

with

$$\Omega_c = \frac{q_1 q_2}{2\pi\epsilon_0 \sqrt{m_1 m_2} \omega_1 \omega_2} \frac{1}{r^3} \quad (4)$$

Here q_i and m_i refer to the charge and mass of the particles, Ω_c is the coupling strength, and $\Delta z_i = \sqrt{\frac{\hbar}{2m_i\omega_i}}(a_i + a_i^\dagger)$ denotes the vibrational amplitude of the motion of ion i ; \hbar is $h/2\pi$ (where h is Planck’s constant). The quantum-mechanical creation and annihilation operators, a_i and a_i^\dagger , act on the individual harmonic oscillators with frequencies ω_i . Rapidly oscillating terms have been neglected in equation (3).

At resonance, that is, for $\omega_1 = \omega_2$, the coupling described by equation (3) leads to a complete exchange of motional states between the two ions after time $T_{\text{swap}} = \pi/\Omega_c$. This is analogous to two coupled

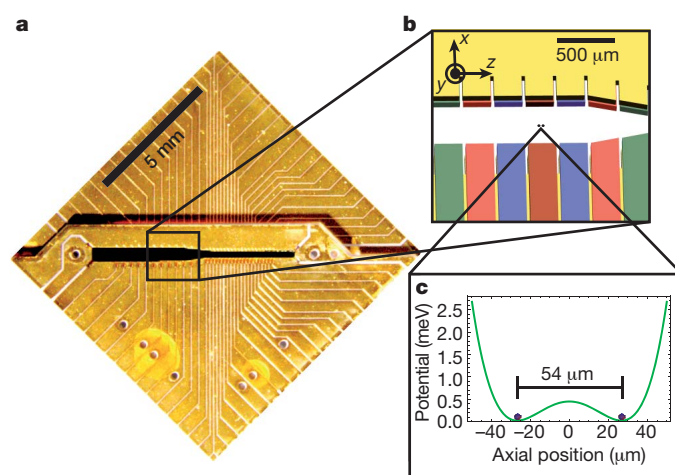


Figure 1 | Creation of trap potential. **a**, Image of the two-layer segmented trap with gold-on-alumina electrodes. The electrodes are 250 μm wide with a pitch of 280 μm , and the ion–electrode spacing is 258 μm . **b**, Schematic close-up of the zone of interest. The electrode pairs used to generate the double-well potential are colour-coded: red for positive voltages, blue for negative voltages and green for compensation voltages. **c**, The axial potential calculated with the parameters determined from the motional spectra in Fig. 4.

¹Institut für Experimentalphysik, Universität Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria. ²Institut für Quantenoptik und Quanteninformation der Österreichischen Akademie der Wissenschaften, Technikerstraße 21a, A-6020 Innsbruck, Austria. [†]Present address: Menlo Systems GmbH, Am Klopferstspitz 19a, D-82152, Martinsried, Germany.

pendula connected by a (massless) spring. If one pendulum initially oscillates while the other is at rest, the motion is periodically exchanged between them. The first pendulum comes to a complete stop after a characteristic time, T_{swap} , which, for small coupling, is inversely proportional to the associated spring constant. A quantum-mechanical description may be given using the vibrational quanta, n_i , often labelled ‘phonons’, in the individual wells. Under the dipole–dipole interaction, an initial motional state $|n_1, n_2\rangle$ becomes entangled with all other motional states $|n'_1, n'_2\rangle$ for which $n'_1 + n'_2 = n_1 + n_2$. Only at odd (even) multiples of T_{swap} is the swapped (original) basis state recovered. Notably, the initial state $|0, 1\rangle$ evolves into the Bell state $(|0, 1\rangle + |1, 0\rangle)/\sqrt{2}$ after time $T_{\text{swap}}/2$, yielding a maximally entangled state of motion. This motional entanglement can be mapped on to the internal electronic state of the ions¹⁸.

In the experiment presented here, the coherent energy exchange is demonstrated between singly charged $^{40}\text{Ca}^+$ ions. They are held in an ion trap with gold-on-alumina electrodes arranged in a two-layer geometry (Fig. 1) similar to the one described in ref. 19. Applying d.c. voltages of up to 110 V to seven adjacent electrode pairs, a double-well potential with a trap separation of 54 μm and an axial trap frequency, $f_0 = 537$ Hz, is created (Methods). The ions are Doppler-cooled on the $S_{1/2}$ – $P_{1/2}$ transition at 397 nm using a single, elliptically shaped laser beam, and detected by collecting the fluorescence light on an electron-multiplied CCD (charge-coupled device) camera and on a photomultiplier tube. Two 729-nm laser beams, individually focused on the two trapping sites, are used to perform sideband cooling on the $S_{1/2}$ – $D_{5/2}$ transition and to map out the sideband spectrum²⁰. During the 4-ms period of sideband cooling, the trapping sites are alternately illuminated with 729-nm light. This alternation is carried out at the approximate rate of the energy exchange to achieve an imbalance of phonon population between the traps. By comparing Rabi-oscillations on the red and blue sideband²¹, the mean phonon number in the first well is calculated. Figure 2 shows the oscillatory behaviour of the phonon number in this well as a function of waiting time. The theoretical fit to the data (Methods) indicates an exchange within $T_{\text{swap}} = 222(10)$ μs and an initial phonon population of $\langle n_2 \rangle = 9(1)$ in the second well; all numbers in parentheses denote 1σ standard deviations with respect to the last digit. Currently, the ions experience a heating rate, $\langle \dot{n} \rangle$, of 1.3(7) quanta per millisecond, which limits the efficiency of the sideband

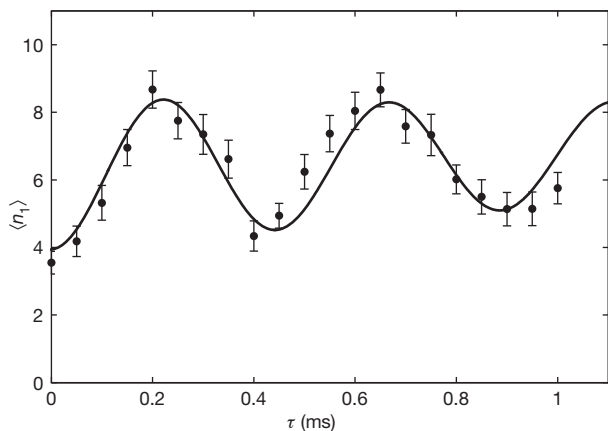


Figure 2 | Energy exchange between two trapped ions over a distance of 54 μm . The data show the average number of phonons in the first trapping well, $\langle n_1 \rangle$, after sideband cooling and a variable waiting time, τ . The sequence for sideband cooling has been arranged to yield different phonon numbers in the two wells, revealing the state exchange as an oscillation of phonon numbers at the level of a few quanta. The observed time for a complete exchange is $T_{\text{swap}} = 222(10)$ μs , indicating a mode splitting of $\Omega_c \approx 2\pi \times 2.25(4)$ kHz. A damping constant $\tau_{\text{damp}} = 3(2)$ ms and a constant background heating, $\langle \dot{n} \rangle$, of 1.3(7) quanta per millisecond are inferred from the fit to the data. The error bars indicate one standard deviation as inferred from Monte Carlo simulations. Lateral deviations of the data from the fit are attributed to small drifts in the resonance condition which modify the exchange rate.

cooling and is comparable to other room-temperature traps of similar size²².

Ideally, the exchange rate would be significantly larger than the average heating rate. To enhance the exchange rate, strings of several ions can be used in the trapping wells. Equation (4) suggests that the coupling strength scales in proportion to the number of ions. This assumes the ion strings to be point-like objects. In practice, the extended nature of the strings can cause a significantly faster increase in the coupling strength.

This increase is demonstrated by mapping out the ions' excitation spectrum as the trapping frequencies of the individual sites are scanned through the resonance condition. The frequency scan is achieved by applying a control voltage U_{ax} to an outer trap-electrode pair (Methods). The dipole–dipole coupling manifests itself as an avoided crossing, separating the mode frequencies by $\Omega_c/2\pi$. Close to resonance, the motion of the ion strings is strongly coupled, and the oscillation can be excited with 729-nm light on either of the two trapping sites²⁰. Figure 3a shows an example of this avoided crossing measured with two ions in each well, while Fig. 3b represents an individual sideband spectrum. The motional spectra of five ion configurations have been analysed, using up to three ions in each well. The configurations and corresponding mode spectra are displayed in Fig. 4. The curves present calculations using a common set of fit parameters for the potential and for the action of the control voltage U_{ax} (Methods), and explicitly include the extended nature of the ion strings. For the configuration with one ion per well, the observed splitting of 1.9 kHz agrees within one standard deviation with the energy-exchange rate from Fig. 2, which was measured with the same trap parameters. The data further show that the splitting is increased from 1.9(3) kHz to 14(1) kHz by using up to three ions in each well, without any modification of the external potential. The sevenfold increase of the coupling is beyond the factor of three that is expected from a simple point-charge model and is due to the anharmonicity of the individual potential wells: as the outer potential walls are steeper than the inner ones, the ion-strings' centres of mass get closer as more ions are added. At the same time, the average oscillation frequency is reduced. Both of these effects lead to an increase in the coupling strength (equation (4)). The extent of the ion strings provides a comparable increase, as it is non-negligible with respect to the inter-well distance.

The demonstrated coupling can be used in diverse schemes to create entanglement or to perform gates. (See also independent experiments at NIST Boulder with trapped $^9\text{Be}^+$ ions²³.) The creation of Bell states

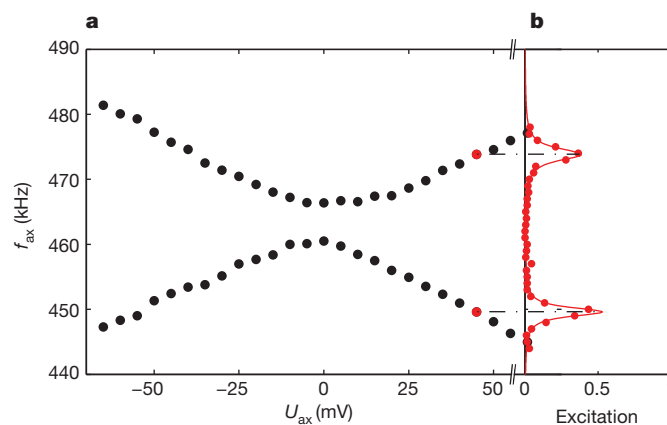


Figure 3 | Demonstration of avoided crossing derived from sideband spectra. **a**, Oscillation frequencies of four trapped ions (two in each well) as a function of the axial control voltage U_{ax} , yielding a mode splitting of 5.5(3) kHz. The data points (filled circles) correspond to peaks in individual sideband spectra taken on the $S_{1/2}$ – $D_{5/2}$ transition. The error bars are smaller than the filled circles. **b**, Example of an individual sideband spectrum. The resonant frequencies are marked with a dot-dashed line, and the corresponding data points in **a** are coloured red.

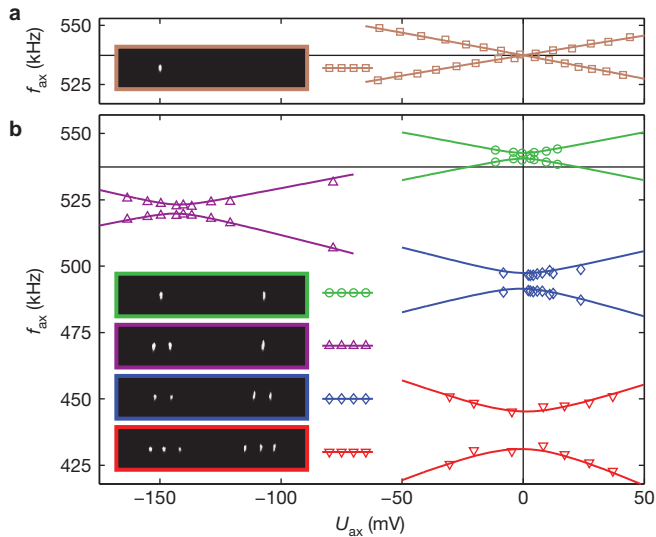


Figure 4 | Experimentally observed dipole-dipole coupling for various ion configurations in a double-well potential. The graphs display the oscillation frequencies of the two lowest vibrational modes as a function of control voltage U_{ax} , and reveal the mode splitting at resonance; the horizontal line in each panel shows f_0 . **a**, Unmodified trap frequencies, as measured with only a single ion in the left well (negative slope) or in the right well (positive slope). **b**, Spectra with up to three ions in each well, as depicted in the insets (inset border colour matches data symbol and line colour). The lines represent predictions from numerical calculations and fit all data simultaneously. A small drift in the control voltage of 7 mV over one hour has been taken into account, aligning the spectra on top of one another. The mode splittings (kHz) for the four different configurations are 1.9(3), 2.4(9), 6.0(3) and 14(1). The shift to lower voltages for three ions is due to the asymmetry in their configuration, (2+1). The vertical elongation of the ion pictures is due to aberrations arising from an off-axis position of the ions relative to the imaging axis.

is discussed above and requires ground-state cooling of the ions. One route to perform a thermally robust quantum gate has been described¹¹, but it requires a state-dependent pushing force by an appropriate laser. It should be noted that a fast variant²⁴ of the thermally robust Mølmer-Sørensen gate²⁵ may also be implemented in the system under consideration: if the required bichromatic laser beams are tuned to the centre of the mode splitting, a gate can be performed within $T_{\text{gate}} = 4\pi/\Omega_c$.

Our coupling scheme also has implications for possible architectures for quantum computation. It is straightforward to imagine a linear array of ion traps, each holding a string of ions, able to interact axially. One could further imagine a two-dimensional array of traps in which parallel dipoles can also be coupled. Bringing rows of traps pairwise into resonance allows the creation of linear cluster states within two steps. A two-dimensional cluster state would only require two further steps by pair-wise coupling of columns²⁶.

The amplification of the dipole-dipole interaction is not limited to ions; the presented technique may be directly transferred to the coupling of trapped Rydberg atoms²⁷. Nor is the interaction limited to systems of equal mass, as demonstrated by the mode-splitting for the asymmetric configuration (2+1) (Fig. 4). Atomic and ionic systems may thus be combined. Neutral atoms have already been brought close to ions^{28,29}. When the ions' oscillation frequency is tuned into resonance with adjacent levels of a Rydberg atom, coupling between ions and Rydberg atoms may become achievable.

METHODS SUMMARY

Calculation of eigenfrequencies in the double-well potential. A symmetric double-well potential is created along the trap axis by applying 2.8 V, 110.4 V, −16.8 V, 13.7 V, −16.8 V, 110.4 V and −33.0 V to the seven adjacent electrode pairs marked in Fig. 1b. The control voltage, U_{ax} , is additionally applied to the leftmost electrode pair, creating a nearly homogeneous electric field in the centre

region. This shifts the trap frequencies of the two wells in opposite directions and is used to tune the wells through resonance. In the region of interest the double-well potential is well described by a fourth-order polynomial, where the zeroth order term does not influence the dynamics and the third-order coefficient is eliminated by appropriate choice of origin:

$$E_{\text{pot}}(z) = q(\alpha_1 z + \alpha_2 z^2 + \alpha_4 z^4) \quad (5)$$

The polynomial coefficients, α_2 and α_4 , determine both the uncoupled trap frequency, ω_0 , and the inter-well separation, r , at the symmetry point³⁰ ($\alpha_1 = 0$). The action of the tuning voltage is modelled as a linear contribution to the coefficients, whereby $\alpha_1 = 4.61 \times 10^{-24} \text{ m}^{-1} \times U_{ax}$ and $\alpha_2 = \alpha_2(U_{ax} = 0) + 0.9 \times 10^{-20} \text{ m}^{-2} \times U_{ax}$. The equilibrium positions for N ions, $\mathbf{z}_0^N = (z_{0,1}, \dots, z_{0,N})$, are those minimizing the total energy, E_{tot}^N :

$$E_{\text{tot}}^N(\mathbf{z}^N) = \sum_{i=1}^N E_{\text{pot}}(z_i) + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{q^2}{4\pi\epsilon_0} \times \frac{1}{|z_i - z_j|} \quad (6)$$

The theoretical data for the vibrational mode frequencies, ω_b , in Fig. 4 are computed from the two lowest eigenvalues of the Hessian matrix of $E_{\text{tot}}^N(\mathbf{z}^N)$ at positions \mathbf{z}_0^N .

Energy exchange. The theoretical fit to the data in Fig. 2 is calculated using

$$\langle n_1 \rangle(\tau) = \langle n_1 \rangle(0) + \frac{1}{2} \langle \Delta n \rangle \left(1 - \cos\left(\pi \frac{\tau}{T_{\text{swap}}}\right) \right) \exp\left(-\frac{\tau}{\tau_{\text{damp}}}\right) + \langle \dot{n} \rangle \tau \quad (7)$$

with $\langle n_1 \rangle(0)$ denoting the initial mean phonon number in the first well, and $\langle \Delta n \rangle = \langle n_2 \rangle(0) - \langle n_1 \rangle(0)$ representing the difference between the initial mean phonon numbers in each well. T_{swap} is the time required for a complete exchange of motional states. Damping due to decoherence and any slow fluctuations of the experimental parameters is taken into account with τ_{damp} , and $\langle \dot{n} \rangle$ denotes the background heating rate.

Received 12 November 2010; accepted 4 January 2011.

Published online 23 February 2011.

- Hertz, H. Über die Grundgleichungen der Electrodynamik für bewegte Körper. *Ann. Phys. Chem.* **277**, 369–399 (1890).
- Kuhr, S. et al. Ultrahigh finesse Fabry-Perot superconducting resonator. *Appl. Phys. Lett.* **90**, 164101 (2007).
- Gleyzes, S. et al. Quantum jumps of light recording the birth and death of a photon in a cavity. *Nature* **446**, 297–300 (2007).
- Anetsberger, G. et al. Near-field cavity optomechanics with nanomechanical oscillators. *Nature Phys.* **5**, 909–914 (2009).
- Schliesser, A., Del'Haye, P., Nooshi, N., Vahala, K. J. & Kippenberg, T. J. Radiation pressure cooling of a micromechanical oscillator using dynamical backaction. *Phys. Rev. Lett.* **97**, 243905 (2006).
- Kinoshita, T., Wenger, T. & Weiss, D. S. A quantum Newton's cradle. *Nature* **440**, 900–903 (2006).
- Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75**, 281–324 (2003).
- Monroe, C., Meekhof, D. M., King, B. E. & Wineland, D. J. A “Schrödinger Cat” superposition state of an atom. *Science* **272**, 1131–1136 (1996).
- Jost, J. D. et al. Entangled mechanical oscillators. *Nature* **459**, 683–685 (2009).
- Blinov, B. B., Moehring, D. L., Duan, L.-M. & Monroe, C. Observation of entanglement between a single trapped atom and a single photon. *Nature* **428**, 153–157 (2004).
- Cirac, J. I. & Zoller, P. A scalable quantum computer with ions in an array of microtraps. *Nature* **404**, 579–581 (2000).
- DiVincenzo, D. P. The physical implementation of quantum computation. *Fortschr. Phys.* **48**, 771–783 (2000).
- Kielpinski, D., Monroe, C. & Wineland, D. J. Architecture for a large-scale ion-trap quantum computer. *Nature* **417**, 709–711 (2002).
- Deslauriers, L. et al. Scaling and suppression of anomalous heating in ion traps. *Phys. Rev. Lett.* **97**, 103007 (2006).
- Seidelin, S. et al. Microfabricated surface-electrode ion trap for scalable quantum information processing. *Phys. Rev. Lett.* **96**, 253003 (2006).
- Leibbrandt, D. R. et al. Modeling ion trap thermal noise coherence. *Quant. Inform. Comput.* **7**, 052–072 (2007).
- Labaziewicz, J. et al. Suppression of heating rates in cryogenic surface-electrode ion traps. *Phys. Rev. Lett.* **100**, 013001 (2008).
- Cirac, J. I. & Zoller, P. Quantum computations with cold trapped ions. *Phys. Rev. Lett.* **74**, 4091–4094 (1995).
- Schulz, S. A., Poschinger, U., Ziesel, F. & Schmidt-Kaler, F. Sideband cooling and coherent dynamics in a microchip multi-segmented ion trap. *N. J. Phys.* **10**, 045007 (2008).
- Roos, C. F. et al. Quantum state engineering on an optical transition and decoherence in a Paul trap. *Phys. Rev. Lett.* **83**, 4713–4716 (1999).
- Meekhof, D. M., Monroe, C., King, B. E., Itano, W. M. & Wineland, D. J. Generation of nonclassical motional states of a trapped atom. *Phys. Rev. Lett.* **76**, 1796–1799 (1996).
- Turchette, Q. A. et al. Heating of trapped ions from the quantum ground state. *Phys. Rev. A* **61**, 063418 (2000).
- Brown, K. R. et al. Coupled quantized mechanical oscillators. *Nature* doi:10.1038/nature09721 (this issue).

24. Benhelm, J. *et al.* Towards fault-tolerant quantum computing with trapped ions. *Nature Phys.* **4**, 463–466 (2008).
25. Sørensen, A. & Mølmer, K. Quantum computation with ions in thermal motion. *Phys. Rev. Lett.* **82**, 1971–1974 (1999).
26. Raussendorf, R., Browne, D. E. & Briegel, H. J. Measurement-based quantum computation on cluster states. *Phys. Rev. A* **68**, 022312 (2003).
27. Urban, E. *et al.* Observation of Rydberg blockade between two atoms. *Nature Phys.* **5**, 110–114 (2009).
28. Zipkes, C., Palzer, S., Sias, C. & Köhl, M. A trapped single ion inside a Bose-Einstein condensate. *Nature* **464**, 388–391 (2010).
29. Schmid, S., Härter, A. & Denschlag, J. H. Dynamics of a cold trapped ion in a Bose-Einstein condensate. *Phys. Rev. Lett.* **105**, 133202 (2010).
30. Home, J. P. & Steane, A. M. Electrode configuration for fast separation of trapped ions. *Quant. Inform. Comput.* **6**, 289–325 (2006).

Acknowledgements We thank H. Häffner for discussions at an early state of the project. We acknowledge the support of the EU STREP project MICROTRAP, the Austrian Science Fund (FWF), the EU network SCALA, the European Research Council (ERC) and the Institut für Quanteninformation GmbH.

Author Contributions The experiments were performed by M.H., R.L. and W.H.; M.H., M.B., R.L., W.H. and R.B. contributed to the set-up; the data analysis was performed by M.H. and W.H.; the original idea was devised by W.H. and R.B.; and all authors contributed to the discussion of the results and participated in manuscript preparation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.B. (rainer.blatt@uibk.ac.at).

Fat cells reactivate quiescent neuroblasts via TOR and glial insulin relays in *Drosophila*

Rita Sousa-Nunes¹, Lih Ling Yee¹ & Alex P. Gould¹

Many stem, progenitor and cancer cells undergo periods of mitotic quiescence from which they can be reactivated^{1–5}. The signals triggering entry into and exit from this reversible dormant state are not well understood. In the developing *Drosophila* central nervous system, multipotent self-renewing progenitors called neuroblasts^{6–9} undergo quiescence in a stereotypical spatiotemporal pattern¹⁰. Entry into quiescence is regulated by Hox proteins and an internal neuroblast timer^{11–13}. Exit from quiescence (reactivation) is subject to a nutritional checkpoint requiring dietary amino acids¹⁴. Organ co-cultures also implicate an unidentified signal from an adipose/hepatic-like tissue called the fat body¹⁴. Here we provide *in vivo* evidence that Slimfast amino-acid sensing and Target of rapamycin (TOR) signalling¹⁵ activate a fat-body-derived signal (FDS) required for neuroblast reactivation. Downstream of this signal, Insulin-like receptor signalling and the Phosphatidylinositol 3-kinase (PI3K)/TOR network are required in neuroblasts for exit from quiescence. We demonstrate that nutritionally regulated glial cells provide the source of Insulin-like peptides (ILPs) relevant for timely neuroblast reactivation but not for overall larval growth. Conversely, ILPs secreted into the haemolymph by median neurosecretory cells systemically control organismal size^{16–18} but do not reactivate neuroblasts. *Drosophila* thus contains two segregated ILP pools, one regulating proliferation within the central nervous system and the other controlling tissue growth systemically. Our findings support a model in which amino acids trigger the cell cycle re-entry of neural progenitors via a fat-body–glia–neuroblasts relay. This mechanism indicates that dietary nutrients and remote organs, as well as local niches, are key regulators of transitions in stem-cell behaviour.

In fed larvae, *Drosophila* neuroblasts (Fig. 1a) exit quiescence from the late first instar (L1) stage onwards. This reactivation involves cell enlargement and entry into S phase, monitored in this study using the thymidine analogue 5-ethynyl-2'-deoxyuridine (EdU). Consistent with a previous study¹⁰, we observed that reactivated neuroblast lineages (neuroblasts and their progeny; Fig. 1b) reproducibly incorporated EdU in a characteristic spatiotemporal sequence: central brain → thoracic → abdominal neuromeres (Fig. 1c and Supplementary Fig. 1). Mushroom-body neuroblasts and one ventrolateral neuroblast, however, are known not to undergo quiescence and to continue dividing for several days in the absence of dietary amino acids¹⁴ (Fig. 1a, c, f). This indicates that dietary amino acids are more than mere 'fuel', providing a specific signal that reactivates neuroblasts. However, explanted central nervous systems (CNSs) incubated with amino acids do not undergo neuroblast reactivation unless co-cultured with fat bodies from larvae raised on a diet containing amino acids¹⁴. We therefore tested the *in vivo* requirement for a fat-body-derived signal (FDS) in neuroblast reactivation by blocking vesicular trafficking and thus signalling from this organ using a dominant-negative Shibire dynamin (SHI^{DN}). This strongly reduced neuroblast EdU incorporation, indicating that exit from quiescence *in vivo* requires an FDS (Fig. 1d, e). One candidate we tested was *Ilp6*, known to be expressed by the fat body^{19,20}, but neither fat-body-specific overexpression nor RNA interference of this gene significantly affected neuroblast reactivation (Supplementary

Table 1 and data not shown). Fat-body cells are known to sense amino acids via the cationic amino-acid transporter Slimfast (SLIF), which activates the TOR signalling pathway, in turn leading to the production of a systemic growth signal^{15,21}. We found that fat-body-specific overexpression of the TOR activator Ras homologue enriched in brain (RHEB), or of an activated form of the p110 PI3K catalytic subunit, or of the p60 adaptor subunit, had no significant effect on neuroblast reactivation in fed animals or in larvae raised on a nutrient-restricted diet lacking amino acids (Fig. 1e, f and data not shown). In contrast, global inactivation of *Tor*, fat-body-specific *Slif* knockdown or fat-body-specific expression of the TOR inhibitors Tuberous sclerosis complex 1 and 2 (*Tsc1/2*) all strongly reduced neuroblasts from exiting quiescence (Fig. 1d, e). Together, these results show that a SLIF/TOR-dependent FDS is required for neuroblasts to exit quiescence and that this may be equivalent to the FDS known to regulate larval growth.

Next we investigated the signalling pathways essential within neuroblasts for their reactivation. Nutrient-dependent growth is regulated in many species by the interconnected TOR and PI3K pathways^{22–24} (Supplementary Fig. 2). In fed larvae, we found that neuroblast inactivation of TOR signalling (by overexpression of *TSC1/2*), or PI3K signalling (by overexpression of p60, the Phosphatase and tensin homologue PTEN, the Forkhead box subgroup O transcription factor FOXO or dominant-negative p110), all inhibited reactivation (Fig. 1e). Conversely, stimulation of neuroblast TOR signalling (by overexpression of RHEB) or PI3K signalling (by overexpression of activated p110 or Phosphoinositide-dependent kinase 1 (PDK1)) triggered precocious exit from quiescence (Fig. 1e). RHEB overexpression had a particularly early effect, preventing some neuroblasts from undergoing quiescence even in newly hatched larvae (Supplementary Fig. 3). Hence, TOR/PI3K signalling in neuroblasts is required to trigger their timely exit from quiescence. Importantly, neuroblast overexpression of RHEB or activated p110 in nutrient-restricted larvae, which lack FDS activity¹⁴, was sufficient to bypass the block to neuroblast reactivation (Fig. 1f). Notably, both genetic manipulations were even sufficient to reactivate neuroblasts in explanted CNSs, cultured without fat body or any other tissue (Fig. 1g). Together with the previous results this indicates that neuroblast TOR/PI3K signalling lies downstream of the amino-acid-dependent FDS during exit from quiescence.

To identify the mechanism bridging the FDS with neuroblast TOR/PI3K signalling, we tested the role of the Insulin-like receptor (InR) in neuroblasts (Supplementary Fig. 2). Importantly, a dominant-negative InR inhibited neuroblast reactivation, whereas an activated form stimulated premature exit from quiescence (Fig. 1e). Furthermore, InR activation was sufficient to bypass the nutrient restriction block to neuroblast reactivation (Fig. 1f). This indicates that at least one of the potential InR ligands, the seven ILPs, may be the neuroblast reactivating signal(s). By testing various combinations of targeted *Ilp* null alleles²⁵ and genomic *Ilp* deficiencies^{25,26}, we found that neuroblast reactivation was moderately delayed in larvae deficient for both *Ilp2* and *Ilp3* (*Df(3L)Ilp2–3*) or lacking *Ilp6* activity (Fig. 2a). Stronger delays, as severe as those observed in *InR*^{3f} mutants, were observed in larvae simultaneously lacking the activities of *Ilp2*, 3 and 5 (*Df(3L)Ilp2–3, Ilp5*) or *Ilp1–5* (*Df(3L)Ilp1–5*)

¹Division of Developmental Neurobiology, Medical Research Council National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK.

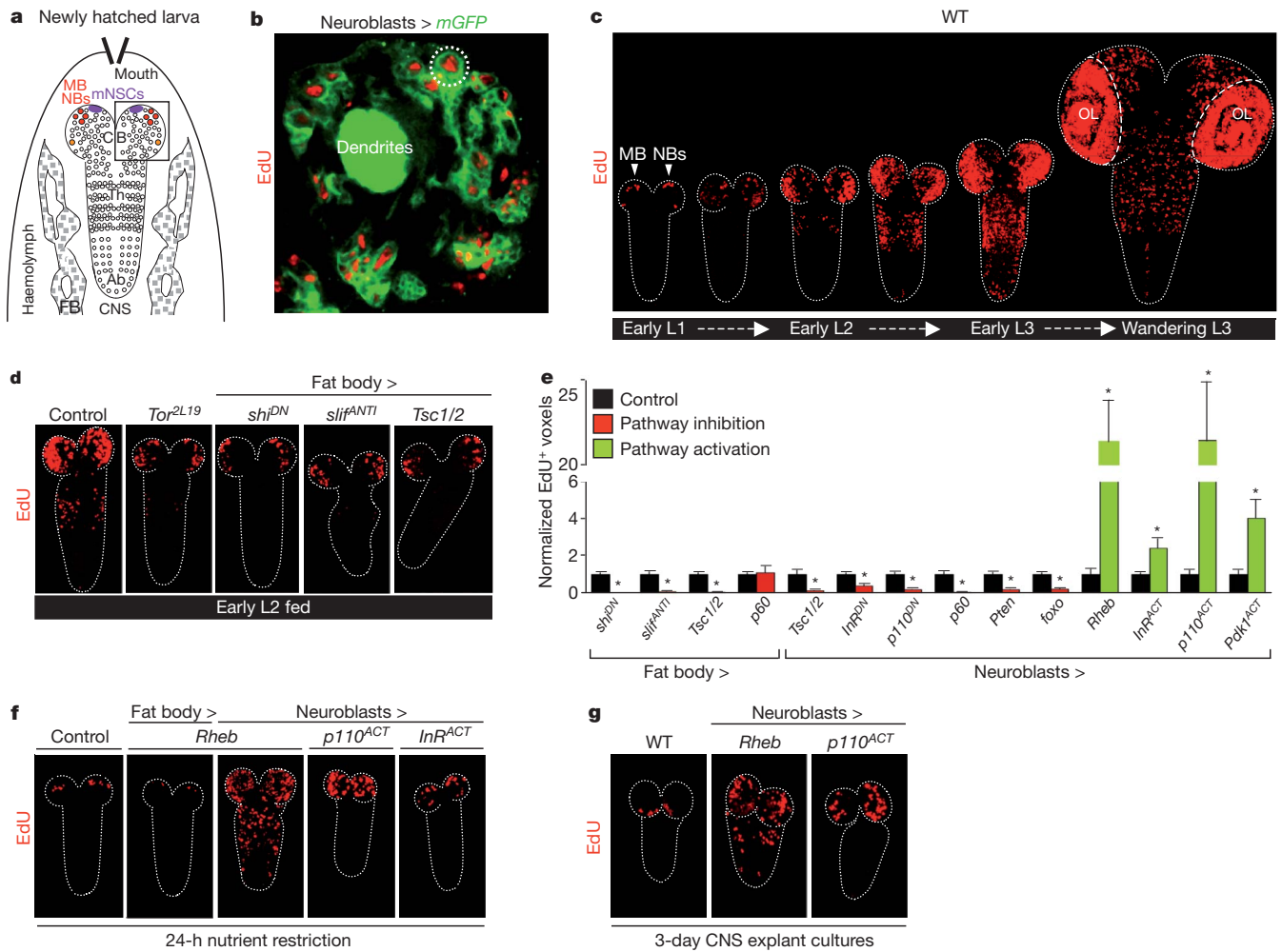


Figure 1 | TOR/PI3K signalling in fat body and neuroblasts regulates reactivation. **a**, Diagram depicting larval fat body (FB) and CNS with central brain (CB), thoracic (Th) and abdominal (Ab) neuromeres, mNSCs, mushroom-body neuroblasts (MB NBs) and other neuroblasts (circles) indicated. **b**, Brain lobe (inset in Fig. 1a), showing EdU incorporation in postembryonic neuroblasts (large cells; for example, dotted circle) and their progeny (smaller cells), labelled with *nab-GAL4* driving membrane GFP

(Fig. 2a). Despite the developmental delay in *Df(3L)Ilp1–5* homozygotes^{25,26}, neuroblast reactivation eventually begins in the normal spatial pattern—albeit heterochronically—in larvae with L3 morphology (Fig. 2b, compare timeline with Fig. 1c). Together, the genetic analysis shows that *Ilp2*, 3, 5 and 6 regulate the timing but not the spatial pattern of neuroblast exit from quiescence. However, as removal of some ILPs can induce compensatory regulation of others²⁵, the relative importance of each cannot be assessed from loss-of-function studies alone.

Brain median neurosecretory cells (mNSCs) (Fig. 1a) are an important source of ILPs, secreted into the haemolymph in an FDS-dependent manner to regulate larval growth^{16–18,21}. They express *Ilp1*, 2, 3 and 5, although not all during the same development stages^{16–18}. However, we found that none of the seven ILPs could reactivate neuroblasts during nutrient restriction when overexpressed in mNSCs (Supplementary Table 2). Similarly, increasing mNSC secretion using the NaChBac sodium channel²¹ or altering mNSC size using PI3K inhibitors/activators, which in turn alters body growth, did not significantly affect neuroblast reactivation under fed conditions (Fig. 2a, c, Supplementary Fig. 1b and L. Y. Cheng and colleagues, manuscript submitted). Surprisingly, therefore, mNSCs are not the relevant ILP source for neuroblast reactivation. Nonetheless, *Ilp3* and *Ilp6* messenger RNAs were detected in the CNS cortex, at the early L2 stage, in a domain

(Neuroblasts > *mGFP*). **c**, EdU incorporation time course from first-instar (L1) to third-instar (L3) larval stages in the wild-type (WT) CNS. OL, optic lobe. **d**, **f**, **g**, EdU-labelled CNSs from larvae expressing TOR/PI3K components driven by *Cg-GAL4* (Fat body >) or *nab-GAL4* (Neuroblasts >). **e**, Histograms of EdU⁺ voxels from thoracic CNSs of fed larvae, normalized to controls. In this and all subsequent figures, error bars are s.e.m.; **P* < 0.05. See text, Methods and Supplementary Fig. 2 for details of molecules expressed.

distinct from the *Ilp2*⁺ mNSCs (Supplementary Fig. 4). Two different *Ilp3-lacZ* transgenes¹⁷ indicate that *Ilp3* is expressed in some glia (Repo⁺ cells) and neurons (Elav⁺ cells). An *Ilp6-GAL4* insertion (see Methods) indicates that *Ilp6* is also expressed in glia, including the cortex glia surrounding neuroblasts and the glia of the blood-brain barrier (BBB) (Fig. 3a).

We next assessed the ability of each of the seven ILPs to reactivate neuroblasts when overexpressed in glia or in neurons (Supplementary Table 2). Pan-glial or pan-neuronal overexpression of ILP4, 5 or 6 led to precocious reactivation under fed conditions (Fig. 3b, c). Each of these manipulations also bypassed the nutrient restriction block to neuroblast reactivation, as did overexpression of ILP2 in glia or in neurons, or ILP3 in neurons (Fig. 3d and Supplementary Table 2). In all of these ILP overexpressions, and even when ILP6 was expressed in the posterior *Ultrabithorax* domain (Fig. 3e), the temporal rather than the spatial pattern of reactivation was affected. Importantly, experiments blocking cell signalling with SHI^{DN} indicate that glia rather than neurons are critical for neuroblast reactivation (Fig. 4a, b). Interestingly, glial-specific overexpression of ILP3–6 did not significantly alter larval mass (Fig. 2c). Thus, in contrast to mNSC-derived ILPs, glial-derived ILPs promote CNS growth without affecting body growth.

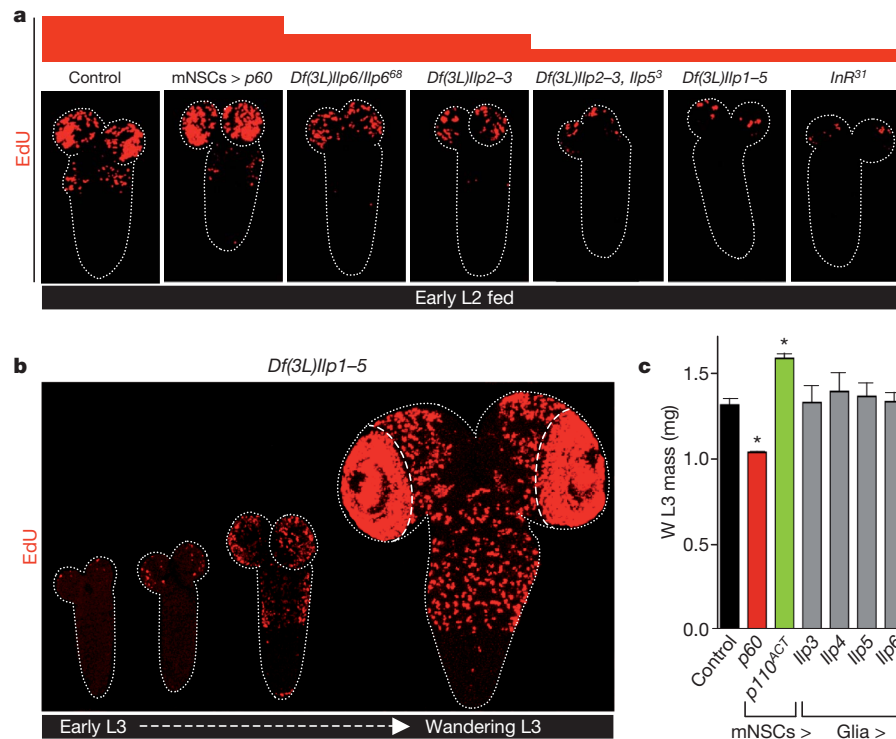


Figure 2 | Insulin-like peptides but not mNSCs control neuroblast reactivation. **a**, EdU-labelled CNSs from various *Ilp* or *InR* mutants show decreased reactivation whereas larvae with *Ilp2-GAL4* driving *UAS-p60* (mNSCs > *p60*) do not. **b**, EdU incorporation time course in the CNS of

Df(3L)Ilp1-5 larvae. **c**, The mass of fed L3 larvae at the wandering (W) stage is significantly altered by *Ilp2-GAL4* (mNSCs >) driving PI3K signalling components but not by *repo-GAL4* (Glia >) driving *Ilp* genes.

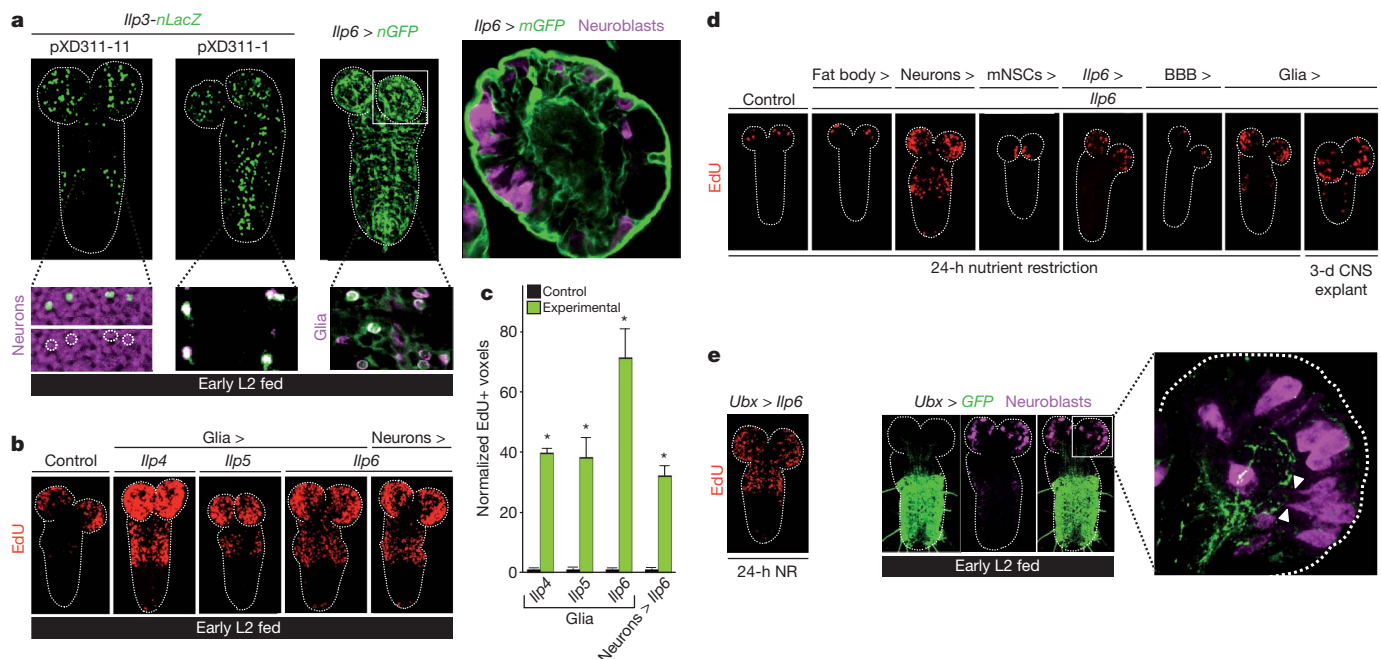


Figure 3 | CNS-specific Insulin-like peptides are sufficient for neuroblast reactivation. **a**, Panels show expression of *Ilp3-nLacZ* in subsets of neurons (XD311-11) and glia (XD311-1) and *Ilp6-GAL4* (*Ilp6 > nGFP* and *Ilp6 > mGFP*) in glia, including BBB and cortex glia. **b**, **d**, EdU-labelled CNSs from larvae overexpressing *Ilp* genes in various cell types (see Methods for *GAL4* drivers used). **c**, Histograms of normalized EdU⁺ voxels in the thoracic CNS for the genotypes in **b**. **e**, *Ilp6* overexpression in the *Ultrabithorax* domain

(*Ubx > Ilp6*) reactivates neuroblasts in the normal spatial pattern during nutrient restriction (NR; left panel). Quiescent/enlarging neuroblasts in the central brain, far from the posterior *Ubx* domain (middle panels), extend cytoplasmic processes (arrowheads) towards the neuropil, close to long *Ubx*⁺ cell processes (right panel). The range of *Ilp6* activity is difficult to determine from this experiment. Neurons, glia and neuroblasts are marked by *Elav*, *Repo* and *Miranda* respectively.

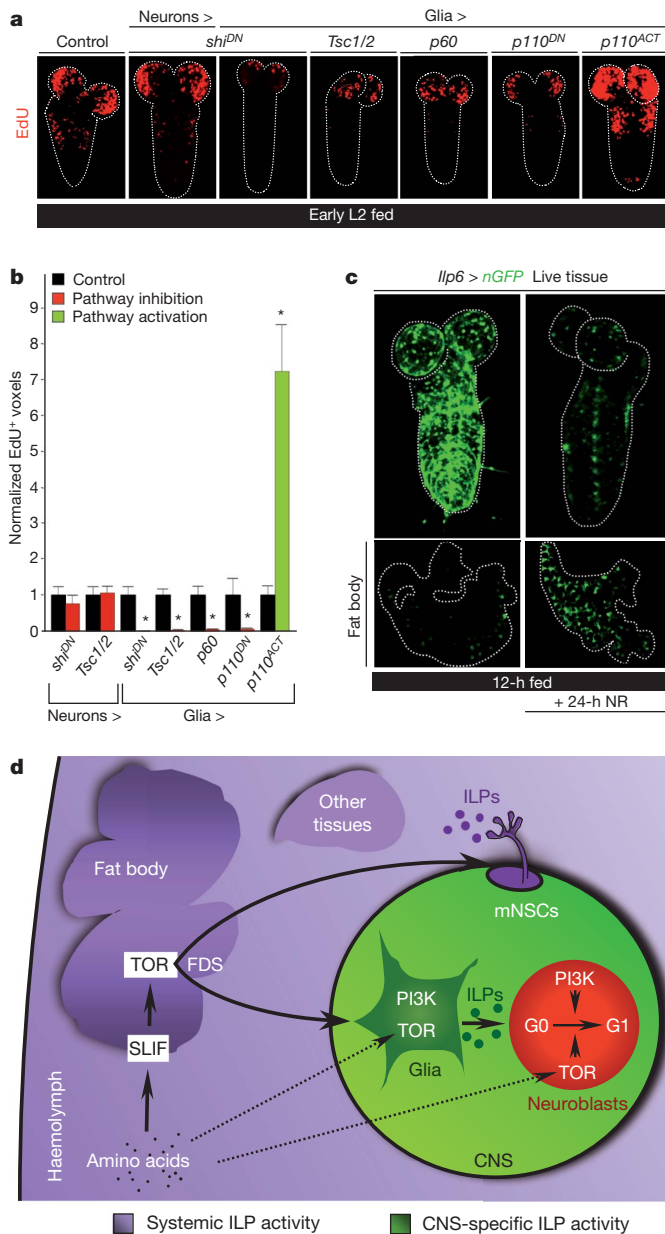


Figure 4 | *Ilp6*-expressing glia are nutritionally regulated. **a**, EdU-labelled CNSs from larvae expressing the components indicated. **b**, Histograms of normalized EdU⁺ voxels in the thoracic CNS for genotypes in **a**. **c**, *Ilp6 > nGFP* expression in the CNS and fat body of fed versus nutrient-restricted (NR) larvae. **d**, Relay model for amino-acid-dependent fat body regulation of CNS and body growth. CNS-restricted (green) and systemic (purple) pools of Insulin-like peptides (ILPs) are functionally segregated. Direct amino-acid sensing by glia and neuroblasts may contribute to neuroblast reactivation (dashed arrows). See text for details.

Focusing on ILP6, we used CNS explant cultures to demonstrate directly that glial overexpression was sufficient to substitute for the FDS during neuroblast exit from quiescence (Fig. 3d). *In vivo*, ILP6 was sufficient to induce reactivation during nutrient restriction when overexpressed via its own promoter or specifically in cortex glia but not in the subperineurial BBB glia, nor in many other CNS cells that we tested (Fig. 3d and Supplementary Table 1). Hence, cortex glia possess the appropriate processing machinery and/or location to deliver reactivating ILP6 to neuroblasts. *Ilp6* mRNA is known to be up- rather than downregulated in the larval fat body during starvation¹⁹ and, accordingly, *Ilp6*-GAL4 activity is increased in this tissue after nutrient restriction (Fig. 4c). Conversely, we found that *Ilp6*-GAL4 is strongly

downregulated in CNS glia during nutrient restriction (Fig. 4c). Thus, dietary nutrients stimulate glia to express *Ilp6* at the transcriptional level. Consistent with this, an important transducer of nutrient signals, the TOR/PI3K network, is necessary and sufficient in glia (but not in neurons) for neuroblast reactivation (Fig. 4a, b). Together, the genetic and expression analyses indicate that nutritionally regulated glia relay the FDS to quiescent neuroblasts via ILPs.

This study used an integrative physiology approach to identify the relay mechanism regulating a nutritional checkpoint in neural progenitors. A central feature of the fat-body–glia–neuroblasts relay model is that glial insulin signalling bridges the amino-acid/TOR-dependent FDS with InR/PI3K/TOR signalling in neuroblasts (Fig. 4d). The importance of glial ILP signalling during neuroblast reactivation is also underscored by an independent study, published while this work was under revision²⁷. As TOR signalling is also required in neuroblasts and glia, direct amino-acid sensing by these cell types may also impinge upon the linear tissue relay. This would then constitute a feed-forward persistence detector²⁸, ensuring that neuroblasts exit quiescence only if high amino-acid levels are sustained rather than transient. We also showed that the CNS ‘compartment’ in which glial ILPs promote growth is functionally isolated, perhaps by the BBB, from the systemic compartment where mNSC ILPs regulate the growth of other tissues. The existence of two functionally separate ILP pools may explain why bovine insulin cannot reactivate neuroblasts in CNS organ culture¹⁴, despite being able to activate *Drosophila* InR *in vitro*²⁹. Given that insulin/PI3K/TOR signalling components are highly conserved between insects and vertebrates, it will be important to address whether mammalian adipose or hepatic tissues signal to glia and whether or not this involves an insulin/IGF relay to CNS progenitors. In this regard, it is intriguing that brain-specific overexpression of IGF1 can stimulate cell-cycle re-entry of mammalian cortical neural progenitors³⁰, indicating utilization of at least part of the mechanism identified here in *Drosophila*.

METHODS SUMMARY

For GAL4/UAS experiments, *Drosophila* were raised at 29 °C unless otherwise stated. Larvae hatching within a 2 h window were transferred to cornmeal food (5.9% glucose, 6.6% cornmeal, 1.2% baker’s yeast, 0.7% agar in water) or nutrient-restricted medium (5% sucrose, 1% agar in PBS) and further synchronized by selecting L2 larvae morphologically from an L1/L2 moulting population. For EdU experiments, dissected CNSs were incubated for 1 h in 10 μM EdU/PBS, fixed for 15 min in 4% formaldehyde/PBS and Alexa Fluor azide was detected according to instructions (Click-iT EdU Imaging Kit, Invitrogen). CNS explants were cultured on 8-μm pore-size inserts in Schneider’s medium, 10% fetal calf serum, 2 mM L-glutamine (Gibco) and 1× Pen Strep (Gibco) in 24-well Transwell plates (Costar) in a humidified chamber at 25 °C. For EdU quantifications, the ‘thoracic’ region used corresponds to the ventral nerve cord from the level of the brain lobes down to A1/A2. EdU⁺ voxels were quantified using Velocity (Improvision) from an average of ten CNSs per experimental genotype, normalized to controls processed in parallel (siblings or half-siblings), using Leica SP5 scans (LAS AF software) with a 1.5-μm-step z-series. For larval mass measurements, triplicates of ~50 wandering L3 male larvae per genotype were transferred to pre-weighed microfuge tubes and wet weights determined using a Precisa XB 120A balance. For all histograms, error bars represent the s.e.m. and *P* values are from two-tailed Student’s *t*-tests with equal sample variance. Further details can be found in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 16 June 2010; accepted 24 January 2011.

Published online 23 February 2011.

1. Dhawan, J. & Rando, T. A. Stem cells in postnatal myogenesis: molecular mechanisms of satellite cell quiescence, activation and replenishment. *Trends Cell Biol.* **15**, 666–673 (2005).
2. Collier, H. A. What’s taking so long? S-phase entry from quiescence versus proliferation. *Nature Rev. Mol. Cell Biol.* **8**, 667–670 (2007).
3. Yanagida, M. Cellular quiescence: are controlling genes conserved? *Trends Cell Biol.* **19**, 705–715 (2009).
4. Chen, E. & Finkel, T. The tortoise, the hare, and the FoxO. *Cell Stem Cell* **5**, 451–452 (2009).

5. Sánchez-García, I., Vicente-Duenas, C. & Cobaleda, C. The theoretical basis of cancer-stem-cell-based therapeutics of cancer: can it be put into practice? *Bioessays* **29**, 1269–1280 (2007).
6. Betschinger, J. & Knoblich, J. A. Dare to be different: asymmetric cell division in *Drosophila*, *C. elegans* and vertebrates. *Curr. Biol.* **14**, R674–R685 (2004).
7. Egger, B., Chell, J. M. & Brand, A. H. Insights into neural stem cell biology from flies. *Phil. Trans. R. Soc. Lond. B* **363**, 39–56 (2008).
8. Doe, C. Q. Neural stem cells: balancing self-renewal with differentiation. *Development* **135**, 1575–1587 (2008).
9. Sousa-Nunes, R., Cheng, L. Y. & Gould, A. P. Regulating neural proliferation in the *Drosophila* CNS. *Curr. Opin. Neurobiol.* **20**, 50–57 (2010).
10. Truman, J. W. & Bate, M. Spatial and temporal patterns of neurogenesis in the central nervous system of *Drosophila melanogaster*. *Dev. Biol.* **125**, 145–157 (1988).
11. Tsuji, T., Hasegawa, E. & Ishiki, T. Neuroblast entry into quiescence is regulated intrinsically by the combined action of spatial Hox proteins and temporal identity factors. *Development* **135**, 3859–3869 (2008).
12. Kambadur, R. *et al.* Regulation of POU genes by castor and hunchback establishes layered compartments in the *Drosophila* CNS. *Genes Dev.* **12**, 246–260 (1998).
13. Ishiki, T., Pearson, B., Holbrook, S. & Doe, C. Q. *Drosophila* neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell* **106**, 511–521 (2001).
14. Britton, J. S. & Edgar, B. A. Environmental control of the cell cycle in *Drosophila*: nutrition activates mitotic and endoreplicative cells by distinct mechanisms. *Development* **125**, 2149–2158 (1998).
15. Colombani, J. *et al.* A nutrient sensor mechanism controls *Drosophila* growth. *Cell* **114**, 739–749 (2003).
16. Brogiolo, W. *et al.* An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. *Curr. Biol.* **11**, 213–221 (2001).
17. Ikeya, T., Galic, M., Belawat, P., Nairz, K. & Hafen, E. Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in *Drosophila*. *Curr. Biol.* **12**, 1293–1300 (2002).
18. Rulifson, E. J., Kim, S. K. & Nüsse, R. Ablation of insulin-producing neurons in flies: growth and diabetic phenotypes. *Science* **296**, 1118–1120 (2002).
19. Slaidina, M., Delanoue, R., Gronke, S., Partridge, L. & Leopold, P. A *Drosophila* insulin-like peptide promotes growth during nonfeeding states. *Dev. Cell* **17**, 874–884 (2009).
20. Okamoto, N. *et al.* A fat body-derived IGF-like peptide regulates postfeeding growth in *Drosophila*. *Dev. Cell* **17**, 885–891 (2009).
21. Géminard, C., Rulifson, E. J. & Leopold, P. Remote control of insulin secretion by fat cells in *Drosophila*. *Cell Metab.* **10**, 199–207 (2009).
22. Polak, P. & Hall, M. N. mTOR and the control of whole body metabolism. *Curr. Opin. Cell Biol.* **21**, 209–218 (2009).
23. Neufeld, T. P. Body building: regulation of shape and size by PI3K/TOR signaling during development. *Mech. Dev.* **120**, 1283–1296 (2003).
24. Teleman, A. A. Molecular mechanisms of metabolic regulation by insulin in *Drosophila*. *Biochem. J.* **425**, 13–26 (2010).
25. Grönke, S., Clarke, D. F., Broughton, S., Andrews, T. D. & Partridge, L. Molecular evolution and functional characterization of *Drosophila* insulin-like peptides. *PLoS Genet.* **6**, e1000857 (2010).
26. Zhang, H. *et al.* Deletion of *Drosophila* insulin-like peptides causes growth defects and metabolic abnormalities. *Proc. Natl Acad. Sci. USA* **106**, 19617–19622 (2009).
27. Chell, J. M. & Brand, A. H. Nutrition-responsive glia control exit of neural stem cells from quiescence. *Cell* **143**, 1161–1173 (2010).
28. Mangan, S. & Alon, U. Structure and function of the feed-forward loop network motif. *Proc. Natl Acad. Sci. USA* **100**, 11980–11985 (2003).
29. Fernandez-Almonacid, R. & Rosen, O. M. Structure and ligand specificity of the *Drosophila melanogaster* insulin receptor. *Mol. Cell. Biol.* **7**, 2718–2727 (1987).
30. Hodge, R. D., D'Ercole, A. J. & O'Kusky, J. R. Insulin-like growth factor-I accelerates the cell cycle by decreasing G1 phase length and increases cell cycle reentry in the embryonic cerebral cortex. *J. Neurosci.* **24**, 10201–10210 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to A. Brand, S. Cohen, B. Edgar, U. Gaul, E. Hafen, C. Klambt, T. Lee, S. Leever, P. Leopold, F. Matsuzaki, I. Miguel-Aliaga, T. Neufeld, R. Palmer, L. Partridge, L. Pick, E. Sanchez-Herrero, H. Stocker, N. Tapon and T. Xu, and also to the Bloomington stock centre and Kyoto National Institute of Genetics (NIG) for *Drosophila* stocks, antibodies and plasmids. We also acknowledge I. Salecker, J.-P. Vincent, A. Bailey, E. Cinnamon, L. Cheng, R. Makki, A. Matheu, P. Pachnis, P. Serpente and I. Stefana for providing advice, reagents and critical reading of the manuscript. The authors were supported by the Medical Research Council (U117584237).

Author Contributions R.S.-N. and A.P.G. designed the experiments, R.S.-N. and L.L.Y. performed the experiments and R.S.-N. and A.P.G. wrote the manuscript. All authors have read and subscribe to the contents of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.P.G. (agould@nimr.mrc.ac.uk).

METHODS

Rearing and staging of *Drosophila* larvae. To assist larval genotyping, lethal chromosomes were re-established over *Dfd-YFP* balancers. For EdU experiments, crosses were performed in cages with grape-juice plates (25% (v/v) grape-juice, 1.25% (w/v) sucrose, 2.5% (w/v) agar) supplemented with live yeast paste. For GAL4/UAS experiments, larvae hatched within a 2-h time window were transferred to our standard cornmeal food (5.9% w/v glucose, 6.6% cornmeal, 1.2% baker's yeast, 0.7% agar in water) or to nutrient-restricted (NR) medium (5% sucrose, 1% agar in PBS). *Drosophila* were raised at 29 °C throughout, with the following exceptions owing to lethality at high temperature: *tub-GAL80^{ts}*, *repo > shi^{DN}* and *tub-GAL80^{ts}, repo > Ilp2* *Drosophila* were raised at 25 °C during embryogenesis and 29 °C during larval development, other *Ilp2* overexpressions were performed at 25 °C throughout. At the time of dissection, development was further synchronized by selecting L2 animals morphologically from a mixed L1/L2 moulting population. *Df(3L)Ilp2-3*, *Ilp5³* and *Df(3L)Ilp1-5* homozygotes develop considerably slower than controls, so EdU-incorporation experiments used morphological staging after the L1/L2 and L2/L3 moults. Co-expression of *Dcr-2* was used to enhance knockdown efficiency for the *slif^{ANTI}* allele, in which antisense *slif* sequences are under UAS control from an EP element insertion. As absolute numbers of reactivated neuroblasts can vary with small differences in temperature and humidity, parallel control experiments were carried out for each genetic background (using the siblings or half-siblings of experimental animals), rather than using a single control.

***Drosophila* strains.** Stocks used in this study were: *Tor^{2L19}* (ref. 31), *InR³¹* (ref. 32), *Df(3L)Ilp1-5*, *Df(X)Ilp6* and *Df(X)Ilp7* (ref. 33), *Ilp1¹*, *Ilp2¹*, *Ilp3¹*, *Ilp4¹*, *Ilp5¹*, *Ilp5³*, *Df(X)Ilp6⁴¹*, *Df(X)Ilp6⁶⁸*, *Ilp7¹*, *Df(3L)Ilp2-3*, *Df(3L)Ilp1-4* (ref. 34), *slif^{ANTI}* (ref. 35), *UAS-Tsc1/2* (ref. 36), *UAS-Rheb*, *UAS-InR^{DN}* = *UAS-InR^{K1409A}*, *UAS-InR^{ACT}* = *UAS-InR^{A1325D}*, FB driver = *Cg-GAL4* (ref. 37), pan-glia driver = *repo-GAL4* (ref. 38), *OK107-GAL4* (ref. 39), *eg-GAL4* (ref. 40), *DopR-GAL4* (ref. 41), *btl-GAL4* (ref. 42), *UAS-CD8::GFP*, *FRT82B*, *tub-GAL80^{ts}*, *Sco/CyO*, *Dfd-YFP* and *Dr/TM6B*, *Sb*, *Dfd-YFP* (Bloomington *Drosophila* Stock Center), *UAS-p110^{DN}* = *UAS-p110^{A2860C}* and *UAS-p110^{ACT}* = *UAS-p110^{CAAX}* (ref. 43), *UAS-Pten* (ref. 44), *UAS-p60* (ref. 45), *UAS-Pdk^{ACT}* = *UAS-Pdk1^{A467V}* (ref. 46), *UAS-foxo* (ref. 47), *UAS-Dcr2* (VDR), *UAS-Ilp7* (ref. 48), *UAS-Ilp1*, *UAS-Ilp2*, *UAS-Ilp3*, *UAS-Ilp4*, *UAS-Ilp5*, *UAS-Ilp6*, *Ilp3-nLacZ^{XP311-1}* and *Ilp3-nLacZ^{XP311-11}* (both recapitulate endogenous *Ilp3* expression in L3 mNSCs, ref. 49), *UAS-shi^{DN}* (ref. 50), NB driver = *nab-GAL4^{NP4604}* (ref. 51), cortex glia driver = *NP577-GAL4* and ensheathing glia driver = *NP6520-GAL4* (ref. 52), *Ilp6-GAL4* = *NP1079-GAL4* (NIG stock centre), subperineurial BBB glia driver = *moody-GAL4* (ref. 53), midline glia driver = *slit-GAL4* (ref. 54), midline glia/neuronal driver = *sim-GAL4* (ref. 55), mNSC driver = *Ilp2-GAL4* (ref. 56), pan-neuronal driver = *n-syb-GAL4* (ref. 57), *Ubx-GAL4* (ref. 58), *wg-GAL4* (ref. 59), *en-GAL4* was a gift from A. Brand via J.-P. Vincent, *repo-FLP* (ref. 60).

EdU detection, immunostaining, *in situ* hybridization and imaging. L1 and L2 tissues were immobilized on poly-L-Lysine-coated slides for all stainings, except for CNS explants. For EdU experiments, dissected CNSs were incubated for 1 h in 10 µM EdU/PBS, fixed for 15 min in 4% formaldehyde/PBS, followed by detection of Alexa Fluor azide according to the manufacturer's instructions (Click-iT EdU Imaging Kit, Invitrogen) and washing in 0.1% Triton/PBS. Antibody staining and *in situ* hybridization were performed according to standard protocols. Primary antibodies used in this study were: rabbit anti-β-Galactosidase (Molecular Probes) 1/2,000; rabbit anti-GFP (Invitrogen) 1/1,000; mouse anti-Repo 1/20; mouse anti-Miranda 1/20 and rat anti-Elav 1/100 (Developmental Studies Hybridoma Bank); pre-adsorbed alkaline-phosphatase-conjugated sheep anti-digoxigenin (Roche) 1/2,000. Secondary antibodies used were: F(ab')₂ fragments conjugated to either Alexa-Fluor-488, Alexa-Fluor-633 (Molecular Probes) or Cy3 (Jackson), used at 1/250–1/2,000. Live tissues were photographed in PBS. Fixed tissues labelled for fluorescence microscopy were mounted in Vectashield (Vector Laboratories) whereas those processed for *in situ* hybridization were mounted in 80% glycerol. Fluorescent images were acquired with a Leica SP5 confocal microscope (LAS AF software) and bright-field images were acquired with a Zeiss Axiophot2 microscope (AxioVision software). Images of the whole CNS are projections of a 1.5-µm-step z-series. Images of fat body and of high-magnification double-labels of parts of the CNS are single sections except for the right panel of Fig. 3e, which is a projection of 13 sections from a z-series.

CNS explant cultures. Explanted CNSs from larvae hatched within a 2-h window were cultured for 3–4 days on 8-µm pore-size inserts in 10 µM EdU in Schneider's medium, 10% fetal calf serum, 2 mM L-glutamine (Gibco) and 1 × Pen Strep (Gibco), in 24-well Transwell plates (Costar) placed in a humidified chamber at 25 °C.

Quantification of EdU incorporation. The 'thoracic' region used for EdU quantifications corresponds to the ventral nerve cord from brain-lobe level down to A1/A2, distinguishable from more posterior neuromeres by a sharp transition in neuroblast density (Fig. 1a). The numbers of EdU⁺ voxels per CNS were determined using Velocity (Improvision) from Leica SP5 confocal microscope scans

(LAS AF software) using a 1.5-µm-step z-series. An average of ten CNSs were quantified per experimental genotype and controls (siblings or half-siblings) were processed in parallel. Control and experimental values were normalized using the average number of control EdU⁺ voxels. For all histograms, error bars represent standard error of the mean (s.e.m.) of normalized values and asterisks indicate *P* < 0.05 using two-tailed Student's *t*-tests with equal sample variance.

Larval mass measurements. Wet weights were determined for wandering L3 male larvae, sexed and genotyped in PBS, dabbed dry with tissue and transferred to pre-weighed microfuge tubes. For each data point, triplicate samples, each containing an average of 50 animals per genotype were weighed (Precisa XB 120A balance).

- Oldham, S. *et al.* Genetic and biochemical characterization of dTOR, the *Drosophila* homolog of the target of rapamycin. *Genes Dev.* **14**, 2689–2694 (2000).
- Brogiolo, W. *et al.* An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. *Curr. Biol.* **11**, 213–221 (2001).
- Zhang, H. *et al.* Deletion of *Drosophila* insulin-like peptides causes growth defects and metabolic abnormalities. *Proc. Natl Acad. Sci. USA* **106**, 19617–19622 (2009).
- Gronke, S. *et al.* Molecular evolution and functional characterization of *Drosophila* insulin-like peptides. *PLoS Genet.* **6**, e1000857 (2010).
- Colombani, J. *et al.* A nutrient sensor mechanism controls *Drosophila* growth. *Cell* **114**, 739–749 (2003).
- Tapon, N. *et al.* The *Drosophila* tuberous sclerosis complex gene homologs restrict cell growth and cell proliferation. *Cell* **105**, 345–355 (2001).
- Hennig, K. M., Colombani, J. & Neufeld, T. P. TOR coordinates bulk and targeted endocytosis in the *Drosophila melanogaster* fat body to regulate cell growth. *J. Cell Biol.* **173**, 963–974 (2006).
- Xiong, W. C. *et al.* repo encodes a glial-specific homeo domain protein required in the *Drosophila* nervous system. *Genes Dev.* **8**, 981–994 (1994).
- Connolly, J. B. *et al.* Associative learning disrupted by impaired G_s signaling in *Drosophila* mushroom bodies. *Science* **274**, 2104–2107 (1996).
- Ito, K., Urban, J. & Technau, G. M. Distribution, classification, and development of *Drosophila* glial cells in the late embryonic and early larval ventral nerve cord. *Roux Arch. Dev. Biol.* **204**, 284–307 (1995).
- Hacker, U. *et al.* piggyBac-based insertional mutagenesis in the presence of stably integrated P elements in *Drosophila*. *Proc. Natl Acad. Sci. USA* **100**, 7720–7725 (2003).
- Shiga, Y., Tanaka-Matakatsumi, M. & Hayashi, S. A nuclear GFP/β-galactosidase fusion protein as a marker for morphogenesis in living *Drosophila*. *Dev. Growth Differ.* **38**, 99–106 (1996).
- Leevers, S. J. *et al.* The *Drosophila* phosphoinositide 3-kinase Dp110 promotes cell growth. *EMBO J.* **15**, 6584–6594 (1996).
- Huang, H. *et al.* PTEN affects cell size, cell proliferation and apoptosis during *Drosophila* eye development. *Development* **126**, 5365–5372 (1999).
- Weinkove, D. *et al.* Regulation of imaginal disc cell size, cell number and organ size by *Drosophila* class Ia phosphoinositide 3-kinase and its adaptor. *Curr. Biol.* **9**, 1019–1029 (1999).
- Rintelen, F., Stocker, H., Thomas, G. & Hafen, E. PDK1 regulates growth through Akt and S6K in *Drosophila*. *Proc. Natl Acad. Sci. USA* **98**, 15020–15025 (2001).
- Puig, O., Marr, M. T., Ruhf, M. L. & Tjian, R. Control of cell number by *Drosophila* FOXO: downstream and feedback regulation of the insulin receptor pathway. *Genes Dev.* **17**, 2006–2020 (2003).
- Miguel-Aliaga, I., Thor, S. & Gould, A. P. Postmitotic specification of *Drosophila* insulinergic neurons from pioneer neurons. *PLoS Biol.* **6**, e58 (2008).
- Ikeya, T. *et al.* Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in *Drosophila*. *Curr. Biol.* **12**, 1293–1300 (2002).
- Moline, M. M., Southern, C. & Bejsovec, A. Directionality of wingless protein transport influences epidermal patterning in the *Drosophila* embryo. *Development* **126**, 4375–4384 (1999).
- Maurange, C., Cheng, L. & Gould, A. P. Temporal transcription factors and their targets schedule the end of neural proliferation in *Drosophila*. *Cell* **133**, 891–902 (2008).
- Awasaki, T., Lai, S.-L., Ito, K. & Lee, T. Organization and postembryonic development of glial cells in the adult central brain of *Drosophila*. *J. Neurosci.* **28**, 13742–13753 (2008).
- Schwabe, T. *et al.* GPCR Signaling is required for blood-brain barrier formation in *Drosophila*. *Cell* **123**, 133–144 (2005).
- Scholz, H., Sadlowski, E., Klaes, A. & Klammt, C. Control of midline glia development in the embryonic *Drosophila* CNS. *Mech. Dev.* **64**, 139–151 (1997).
- Kidd, T., Bland, K. S. & Goodman, C. S. Slit is the midline repellent for the Robo receptor in *Drosophila*. *Cell* **96**, 785–794 (1999).
- Rulifson, E. J., Kim, S. K. & Nusse, R. Ablation of insulin-producing neurons in flies: growth and diabetic phenotypes. *Science* **296**, 1118–1120 (2002).
- Pospisilik, J. A. *et al.* *Drosophila* genome-wide obesity screen reveals hedgehog as a determinant of brown versus white adipose cell fate. *Cell* **140**, 148–160 (2010).
- de Navas, L. F., Garaulet, D. L. & Sanchez-Herrero, E. The *Ultrathorax* Hox gene of *Drosophila* controls haltere size by regulating the Dpp pathway. *Development* **133**, 4495–4506 (2006).
- Pfeiffer, S., Alexandre, C., Calleja, M. & Vincent, J. P. The progeny of wingless-expressing cells deliver the signal at a distance in *Drosophila* embryos. *Curr. Biol.* **10**, 321–324 (2000).
- Silies, M. *et al.* Glial cell migration in the eye disc. *J. Neurosci.* **27**, 13130–13139 (2007).

Recapitulation of premature ageing with iPSCs from Hutchinson–Gilford progeria syndrome

Guang-Hui Liu¹, Basam Z. Barkho¹, Sergio Ruiz¹, Dinh Diep², Jing Qu¹, Sheng-Lian Yang¹, Athanasia D. Panopoulos¹, Keiichiro Suzuki¹, Leo Kurian¹, Christopher Walsh¹, James Thompson³, Stephanie Boue⁴, Ho Lim Fung², Ignacio Sancho-Martinez¹, Kun Zhang², John Yates III³ & Juan Carlos Izpisua Belmonte^{1,4}

Hutchinson–Gilford progeria syndrome (HGPS) is a rare and fatal human premature ageing disease^{1–5}, characterized by premature arteriosclerosis and degeneration of vascular smooth muscle cells (SMCs)^{6–8}. HGPS is caused by a single point mutation in the lamin A (*LMNA*) gene, resulting in the generation of progerin, a truncated splicing mutant of lamin A. Accumulation of progerin leads to various ageing-associated nuclear defects including disorganization of nuclear lamina and loss of heterochromatin^{9–12}. Here we report the generation of induced pluripotent stem cells (iPSCs) from fibroblasts obtained from patients with HGPS. HGPS-iPSCs show absence of progerin, and more importantly, lack the nuclear envelope and epigenetic alterations normally associated with premature ageing. Upon differentiation of HGPS-iPSCs, progerin and its ageing-associated phenotypic consequences are restored. Specifically, directed differentiation of HGPS-iPSCs to SMCs leads to the appearance of premature senescence phenotypes associated with vascular ageing. Additionally, our studies identify DNA-dependent protein kinase catalytic subunit (DNAPKcs, also known as PRKDC) as a downstream target of progerin. The absence of nuclear DNAPK holoenzyme correlates with premature as well as physiological ageing. Because progerin also accumulates during physiological ageing^{6,12,13}, our results provide an *in vitro* iPSC-based model to study the pathogenesis of human premature and physiological vascular ageing.

Three HGPS primary fibroblast lines, originally isolated from patients with the classical *LMNA* mutation (Gly608Gly), were transduced with retroviruses encoding OCT4, SOX2, KLF4, c-MYC and green fluorescent protein (GFP). NANOG-positive colonies were effectively obtained when using early passage, but not late passage (>25), HGPS fibroblasts (Supplementary Fig. 1a). We focused on iPSC lines of a well characterized HGPS fibroblast line, AG01972 (refs 9–12). Compared to normal fibroblasts, HGPS fibroblasts (AG01972) showed abnormal nuclear morphology, reduced expression of the lamina components lamin B1 and LAP2β (also known as TMPO), loss of heterochromatin markers H3K9me3, HP1α (also known as CBX5) and HDAC1, and reduced expression of nuclear proliferation marker Ki67 (also known as MKI67; Fig. 1a and Supplementary Fig. 2). From HGPS fibroblasts, we derived six iPSC lines. In addition, we generated control iPSC lines from wild-type fibroblasts (BJ and IMR-90 cell lines). Control and HGPS iPSC lines demonstrated pluripotent gene expression, demethylation of the OCT4 promoter and transgene silencing (Fig. 1b and Supplementary Figs 1, 3 and 4a, and data not shown). They were maintained for more than 50 passages without a loss of pluripotency or the acquisition of detectable morphological or growth abnormalities. The pluripotency of each iPSC line was assessed by differentiation into the three embryonic germ layers *in vitro*, using embryoid body (EB) formation, and/or *in vivo*, by teratoma formation (Supplementary Fig. 5a–d). Out of these lines, we focused primarily on

HGPS-iPSC#4 and BJ-iPSC#3 for our studies (hereafter referred to as HGPS-iPSC or BJ-iPSC). Both BJ-iPSCs and HGPS-iPSCs were able to differentiate towards specialized mesoderm-derivatives such as smooth muscle cells (SMC), endothelial cells (Supplementary Fig. 6a), or beating cardiomyocytes (Supplementary Movies 1 and 2). Moreover, all analysed iPSC lines showed normal chromosomal integrity (Supplementary Fig. 5e). Finally, *LMNA* sequencing confirmed the presence of the classical mutation in HGPS-iPSCs (Supplementary Fig. 5f). Altogether, these data indicate that the somatic cells from HGPS patients, despite their significant premature senescence phenotypes and nuclear defects, have been properly reprogrammed and can be effectively maintained in a pluripotent state.

Lamin A/C protein is expressed in differentiated somatic cells but is absent in embryonic stem cells (ESCs)^{11,14}. Therefore, we next examined the expression of lamin A/C in the generated iPSC lines. As shown in Fig. 2a, lamin A/C expression is significantly downregulated in iPSCs, compared to their parental fibroblasts, whereas lamin B1 transcripts were upregulated. Although progerin should follow a similar pattern of expression as observed for lamin A/C, *LMNA* expression is independent

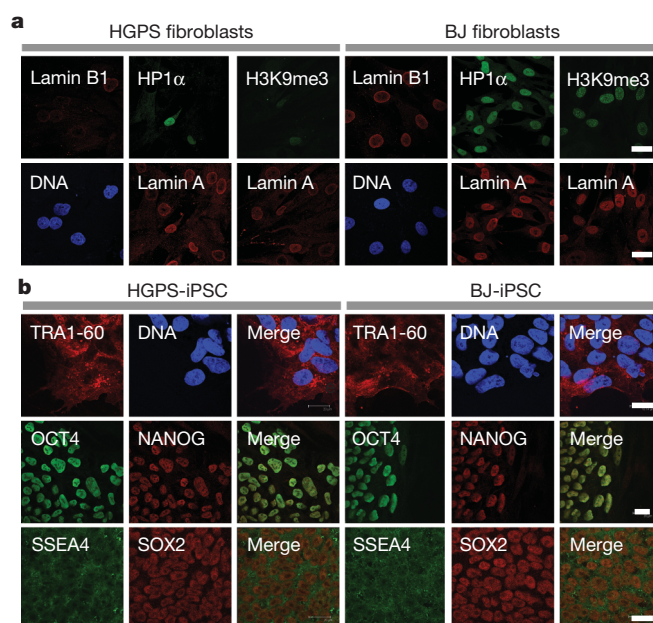


Figure 1 | Generation of iPSCs from HGPS fibroblasts. **a**, Immunofluorescence analysis performed on HGPS (left) and BJ (right) fibroblasts at passage 17 with the indicated antibodies. **b**, Immunofluorescence analysis of the indicated pluripotent markers in HGPS-iPSCs (left) and BJ-iPSCs (right). Nuclei were visualized with Hoechst stain (blue). Scale bar, 20 μm.

¹Gene Expression Laboratory, Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. ²Department of Bioengineering, University of California at San Diego, La Jolla, California 92093, USA. ³Department of Cell Biology, Scripps Research Institute, La Jolla, California 92037, USA. ⁴Center for Regenerative Medicine in Barcelona, Dr. Aiguader 88, 08003 Barcelona, Spain.

of promoter methylation status (Supplementary Fig. 4b)¹⁵. Indeed, a complete loss of progerin mRNA in HGPS-iPSCs was observed (Fig. 2a). Furthermore, expression of lamin A/C and progerin proteins was practically undetectable (Fig. 2b and Supplementary Fig. 4c).

Because HGPS-iPSCs did not express progerin, we examined whether the nuclear abnormalities observed in HGPS fibroblasts would also be absent at the pluripotent stage. Our results indicate that all of the epigenetic, nuclear lamina and proliferation parameters analysed in HGPS-iPSCs are indistinguishable from BJ-iPSCs (Fig. 2b, c and Supplementary Figs 3 and 7). In addition, the nuclei of HGPS-iPSCs displayed the characteristic wrinkles and lobes observed in hESCs and iPSCs (Supplementary Fig. 8), indicative of a reprogramming of the nuclear envelope components. As the nuclear envelope associates with and regulates heterochromatin^{11,16}, we next examined genome-wide CpG methylation in HGPS fibroblasts, BJ fibroblasts, HGPS-iPSCs, BJ-iPSCs, and H9 hESCs. Using bisulphite padlock probes and Illumina sequencing, we captured and quantified the methylation level of an average of 95,932 CpG sites within a set of 16,206 well-annotated differentially methylated regions (DMRs)¹⁷ per cell line (Supplementary Table 1). The correlation coefficient of the global methylation levels between the pluripotent lines (BJ-iPSCs, HGPS-iPSCs and H9 hESCs) and the corresponding fibroblasts indicated that the generated pluripotent lines are much more closely related to each other and to hESCs than the two fibroblast lines (Fig. 2d). Interestingly, 586 autosome genes were found to be associated with regions that showed methylation differences between HGPS and BJ fibroblasts (Supplementary Table 2 and Supplementary Fig. 9a). Furthermore, based on DAVID^{18,19} analysis, we found that these genes were enriched for 21 Gene Ontology terms, most of which were related to development and transcriptional regulation (Supplementary Fig. 9a). In contrast, methylation differences between HGPS-iPSCs and BJ-iPSCs were only found for 33 autosome genes (Supplementary Table 3), which showed no significant functional enrichment. Therefore, the presence of progerin in HGPS fibroblasts seems to lead

to major epigenomic changes in various pathways. These changes were no longer present in HGPS-iPSCs, coinciding with the down-regulation of progerin. Finally, genome-wide mRNA profiling demonstrated that HGPS-iPSCs and BJ-iPSCs are closely related together with H9 hESCs, and different from their parental fibroblasts (Supplementary Fig. 9b–d). These results demonstrate the complete resetting of the nuclear architecture, epigenome and global gene expression in HGPS cells after being reprogrammed to pluripotency.

To test whether the expression of progerin could be re-activated, we first subjected HGPS-iPSCs to *in vitro* differentiation via EB formation. Progerin mRNA was selectively induced in differentiated HGPS-iPSCs, but not in differentiated BJ-iPSCs (Supplementary Fig. 10a). In contrast, lamin A was upregulated in both differentiated HGPS-iPSCs and differentiated BJ-iPSCs (Supplementary Fig. 10a). This reversible suppression of progerin expression by reprogramming, and subsequent reactivation upon differentiation, provides a unique model system to study human premature ageing pathologies. Progerin is known to accumulate mainly in arterial SMCs of HGPS patients, and vascular SMC degeneration is one of the characteristics of HGPS-associated arteriosclerosis^{6,7,20}. In fact, vascular SMC senescence has been involved in the advanced arteriosclerosis of normal populations^{7,21,22}. Therefore, we next asked whether SMCs differentiated from HGPS-iPSCs exhibit premature senescence phenotypes. Using a directed differentiation protocol, we obtained SMC populations from HGPS-iPSCs and BJ-iPSCs, the majority of which expressed characteristic SMC markers such as smooth muscle actin (SMA) and calponin (Supplementary Fig. 6a). Immunoblotting and PCR with reverse transcription (RT-PCR) analyses confirmed the expression of progerin in HGPS-iPSC, but not BJ-iPSC-derived SMCs (hereafter referred to as 'HGPS-SMC' and 'BJ-SMC', Supplementary Fig. 6b, c). To model SMC senescence *in vitro*, the differentiated SMCs were serially passaged in culture. As shown in Fig. 3a–c and Supplementary Fig. 10b, an increasing frequency of misshapen nuclei and a loss of the heterochromatin mark H3K9me3

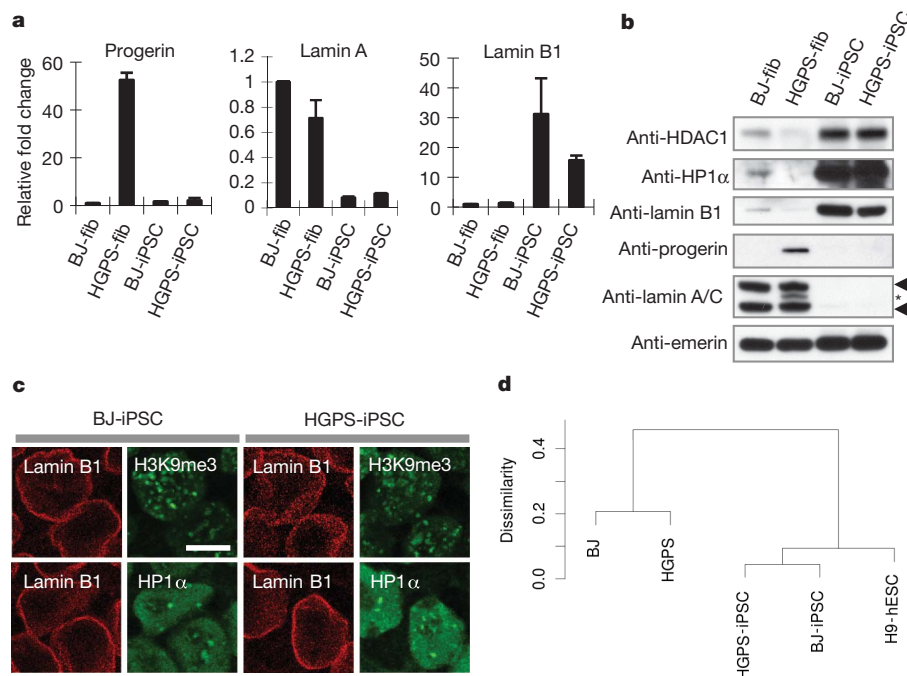


Figure 2 | HGPS-associated nuclear defects are reset in HGPS-iPSCs. **a**, RT-PCR analysis of progerin, lamin A and lamin B1 in the specific cell lines ($n = 3$). BJ- and HGPS-fib, BJ and HGPS fibroblasts. **b**, Immunoblotting analysis of the indicated proteins. Emerin was used as loading control. Asterisk denotes

progerin ($\Delta 50$ lamin A). Arrowheads denote lamin A (top) and lamin C (bottom). **c**, Immunofluorescence analysis performed on BJ-iPSCs and HGPS-iPSCs for detection of the indicated proteins. Scale bar, 10 μ m. **d**, Hierarchical clustering of genome-wide DNA methylation profiles.

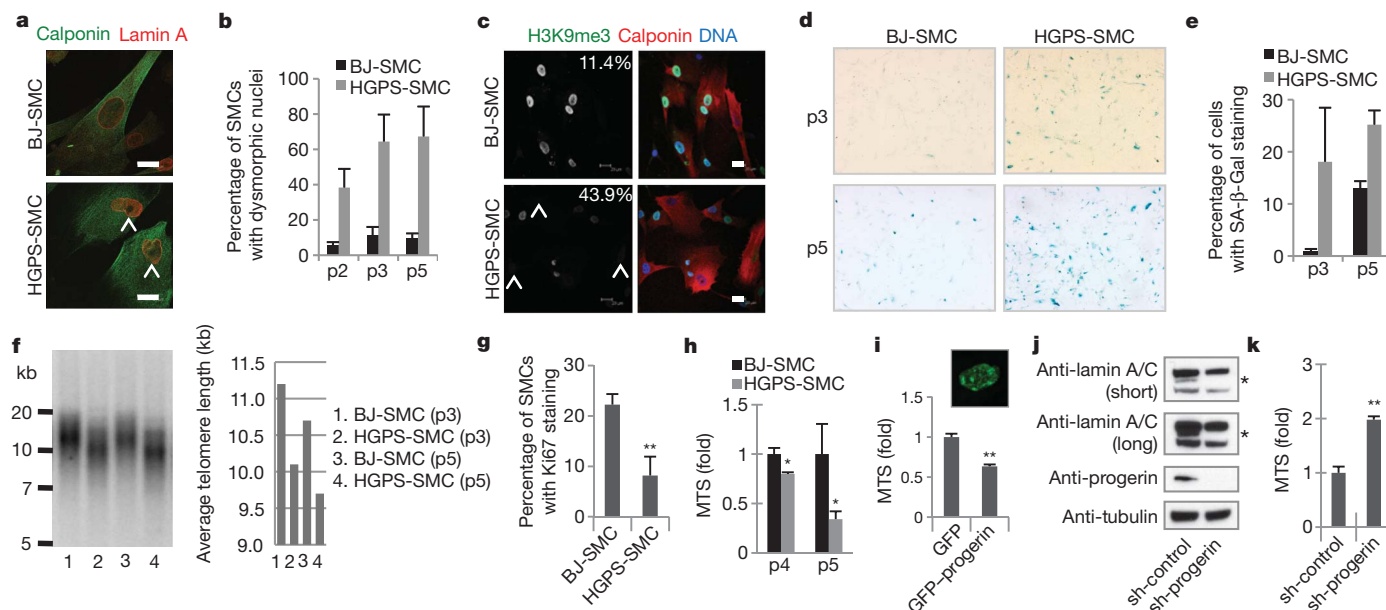


Figure 3 | SMCs expressing progerin show nuclear defects and accelerated senescence. **a**, Immunostaining of calponin and lamin A in iPSC-derived SMCs (p5). Arrowheads denote dysmorphic nuclei. Scale bar, 20 μm. **b**, Percentage of calponin-positive cells showing dysmorphic nuclei, (n = 3, P < 0.001). p2, p3 and p5, passage 2, 3 and 5, respectively. **c**, Immunostaining of H3K9me3 and calponin in iPSC-derived SMCs (p5). Nuclei were visualized with Hoechst stain (blue). Arrowheads denote decreased nuclear H3K9me3 (percentage in corner). Scale bar, 20 μm. **d, e**, Senescence-associated (SA)-β-Gal staining of iPSC-derived SMCs, P < 0.05. **f**, Southern blot analysis of SMCs showing telomere length (left). Quantified average of telomere length (right,

n = 2). **g**, Percentage of Ki67-positive cells in iPSC-derived SMCs (calponin-positive, p3), **P < 0.01. **h, i**, Cell proliferation analysis of iPSC-derived SMCs (n = 3), *P < 0.05 (**h**) or primary vascular SMCs (overexpressing GFP or GFP-progerin, n = 3), **P < 0.01 (**i**). Typical GFP-progerin-positive nucleus showing abnormal morphology (inset). **j**, Immunoblotting of the indicated proteins in shRNA-modified HGPS-iPSCs after 21 days of EB-mediated differentiation. Asterisks denote progerin (Δ50 lamin A). **k**, Cell proliferation analysis of the SMCs derived from shRNA-modified HGPS-iPSCs (p2, n = 3), **P < 0.01.

were specifically observed in HGPS-SMCs after serial passaging. HGPS-SMCs at later passages (that is, passage 5) showed the typical characteristics of premature senescence, including increased senescence-associated-β-galactosidase (SA-β-Gal) staining (Fig. 3d, e and Supplementary Fig. 10c), reduced telomere length (Fig. 3f), a reduced number of Ki67-positive cells (Fig. 3g and Supplementary Fig. 10d), and compromised cell proliferation (Fig. 3h and Supplementary Fig. 10e). We also found a selective upregulation of senescence-related transcripts in HGPS-SMCs (Supplementary Fig. 10f). To test whether the observed HGPS-related cell phenotypes were specific to SMCs, we differentiated HGPS-iPSCs into fibroblasts and measured progerin-associated parameters. Progerin expression in HGPS-iPSC-derived fibroblasts was detectable as early as passage 5 (Supplementary Fig. 11a). However, we were unable to detect a loss of lamina or heterochromatin markers before passage 10 (Supplementary Fig. 11b, c). Nonetheless, these defects were present specifically in HGPS-iPSC-derived fibroblasts, in contrast to control iPSC-derived fibroblasts analysed at similar passage (data not shown). Thus, even though direct comparison of SMCs and fibroblasts is not possible due to their different culture conditions, our observations demonstrate that mesoderm lineages derived from HGPS-iPSCs display a characteristic HGPS phenotype.

We next investigated whether progerin accumulation is the direct cause of the accelerated cell senescence observed in HGPS-SMCs. To this end, we induced ectopic expression of progerin in human primary vascular SMCs. We found that introduction of progerin in wild-type SMCs resulted in compromised cell proliferation and nuclear defects, as we had observed in HGPS-SMCs (Fig. 3i and Supplementary Fig. 12). As a complementary approach, we transduced HGPS-iPSCs with a lentiviral vector expressing a progerin-specific short hairpin RNA (shRNA)²³. The modified iPSCs showed normal karyotypes as well as normal expression of lamina/epigenetic and pluripotent markers

(Supplementary Fig. 13a, b). After EB-based differentiation, both the mRNA and protein levels of progerin, but not those of lamin A, were substantially downregulated in the progerin-shRNA 'corrected' HGPS-iPSCs compared to control cells (Fig. 3j and Supplementary Fig. 13c, d). We next differentiated these 'progerin-free' HGPS-iPSCs into SMCs (Supplementary Fig. 14a). A marked improvement in the proliferation capability, as well as a downregulation of senescence-related transcripts, was found in the SMCs differentiated from the corrected HGPS-iPSCs (Fig. 3k and Supplementary Fig. 14b, c). Furthermore, transduction of progerin shRNA into early passage HGPS-iPSC-derived fibroblasts resulted in a clear restoration of nuclear morphology and heterochromatin markers after extended culture (Supplementary Fig. 15). Taken together, these data identify progerin as the key factor underlying the premature senescence phenotypes of HGPS-iPSC-derived cells.

Because phenotypic characteristics of premature ageing were able to be recapitulated by directed differentiation of the HGPS-iPSCs, we next investigated whether this model could serve to identify novel senescence-related markers. By using a sensitive MudPIT proteomic approach^{24,25}, we identified DNA-dependent protein kinase catalytic subunit (DNAPKcs) as a hitherto unknown binding partner of progerin (Supplementary Table 4). DNAPK holoenzyme, comprising DNAPKcs and its regulatory subunits Ku70/Ku80 (also known as XRCC6 and XRCC5, respectively), is involved in various ageing-related cellular events^{26,27}, and DNAPKcs or Ku70/Ku80-deficient mice exhibit accelerated ageing^{27,28}. To further confirm the association of progerin with DNAPKcs, we performed co-immunoprecipitation experiments. As shown in Fig. 4a, ectopically expressed progerin associated tightly with endogenous DNAPKcs. In contrast, lamin A showed weak interaction with DNAPKcs, whereas both progerin and lamin A exhibited similar binding to lamin B1. Neither progerin nor lamin A co-immunoprecipitated with endogenous WRN protein. Since most of

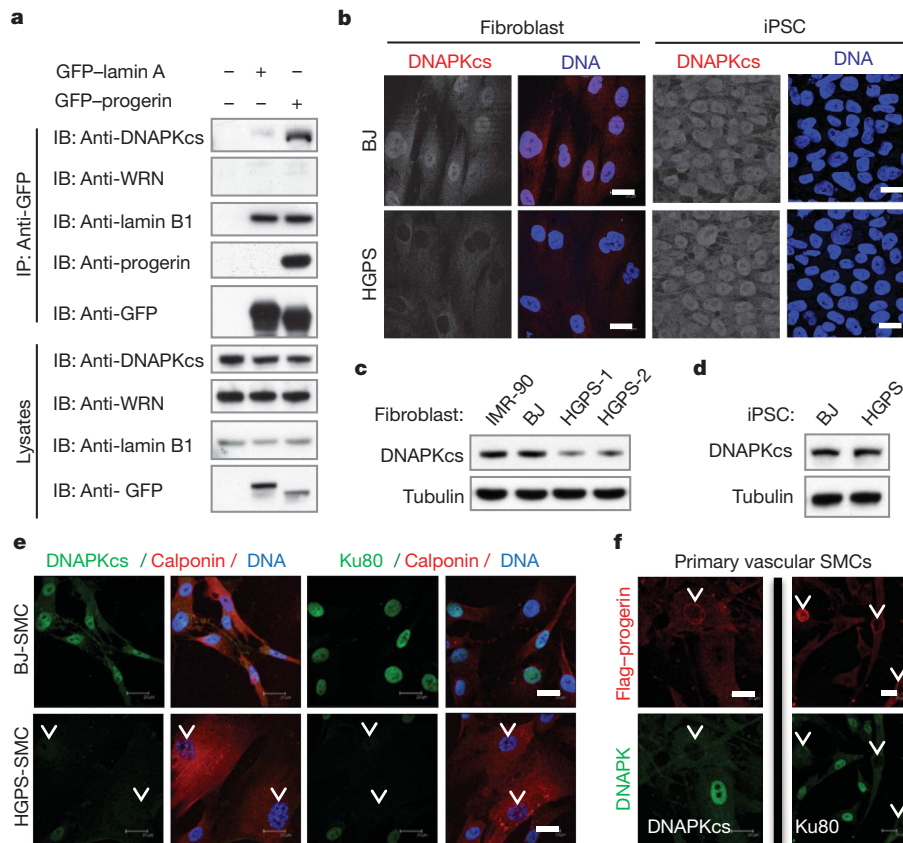


Figure 4 | Decreased expression of DNAPK holoenzyme correlates with premature cell ageing. **a**, Extracts from BJ fibroblasts expressing GFP, GFP-lamin A, or GFP-progerin, were immunoprecipitated (IP) with a GFP antibody and examined by immunoblotting analysis (IB). **b**, DNAPKcs staining in the indicated cell lines. **c**, **d**, Immunoblot analysis of DNAPKcs expression.

the nuclear proteins in complex with lamin A are destabilized in HGPS cells^{9,29}, we analysed the status of DNAPKcs in primary HGPS fibroblasts. We observed decreased nuclear DNAPKcs in HGPS fibroblasts when compared to normal fibroblasts (Fig. 4b, c). In addition, the regulatory subunits Ku70/Ku80 were also downregulated in HGPS fibroblasts (Supplementary Fig. 16a). Interestingly, we detected a complete restoration of DNAPKcs/Ku70/Ku80 expression in HGPS-iPSCs (Fig. 4b, d and Supplementary Figs 3 and 16b), although a deficiency in the expression of these proteins reappeared after differentiation into SMCs (Fig. 4e). These observations indicate that the downregulation of DNAPKcs in HGPS cells is dependent on the accumulation of progerin in differentiated cells. In fact, ectopic expression of progerin in primary vascular SMCs diminished DNAPKcs/Ku80 protein expression (Fig. 4f and Supplementary Fig. 17). We next tested whether inactivation of DNAPKcs could partially contribute to some of the phenotypes observed in HGPS-SMCs. We found that knockdown of DNAPKcs reduced the proliferation of primary vascular SMCs (Supplementary Fig. 18a). Finally, we extended our results and found that progressive loss of DNAPKcs/Ku70/Ku80 also occurs in fibroblasts isolated from normally ageing individuals (Supplementary Fig. 18b). Overall, our data indicate that deficiency of the DNAPK holoenzyme may constitute a novel marker for premature as well as physiological ageing.

In summary, our results not only highlight the plasticity of the lamina-epigenetics axis, but also point to the fact that the altered structure of the nuclear envelope, as well as the epigenetic modifications that accumulate during physiological ageing¹² or under specific disease conditions^{1,11}, can be restored to normalcy by reprogramming (Supplementary Fig. 19). The gradual onset and complexity of ageing has impeded progress in understanding the pathogenesis of ageing-related

e, f, Immunostaining of the indicated proteins in iPSCs-derived SMCs (**e**) or primary vascular SMCs overexpressing progerin (**f**). Arrowheads denote decreased DNAPKcs or Ku80. Nuclei were visualized with Hoechst stain (blue). Scale bar, 20 μ m.

cardiovascular disorders. Recently, striking similarities between normal ageing-associated and HGPS-associated arteriosclerosis have been reported^{6,16}. Indeed, the levels of progerin increase gradually during physiological ageing⁶. Our study provides the first evidence that, in a progerin-dependent manner, HGPS-iPSC-derived SMCs reach senescence-related phenotypes earlier than their normal counterparts. The iPSC-based accelerated ageing model presented here and in ref. 30 may provide an avenue to model and study the pathogenesis of human ageing-related vascular diseases as well as various human laminopathies¹.

METHODS SUMMARY

iPSCs were generated from human fibroblasts with retroviruses encoding OCT4/SOX2/KLF4/c-MYC/GFP, and grown on MEF feeder cells or Matrigel. SMCs were differentiated from iPSCs-derived CD34⁺ progenitor cells following an OP9-based protocol.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 September 2010; accepted 1 February 2011.

Published online 23 February 2011.

1. Worman, H. J., Ostlund, C. & Wang, Y. Diseases of the nuclear envelope. *Cold Spring Harb. Perspect. Biol.* **2**, a000760 (2010).
2. Burtner, C. R. & Kennedy, B. K. Progeria syndromes and ageing: what is the connection? *Nature Rev. Mol. Cell Biol.* **11**, 567–578 (2010).
3. Kudlow, B. A., Kennedy, B. K. & Monnat, R. J. Jr. Werner and Hutchinson–Gilford progeria syndromes: mechanistic basis of human progeroid diseases. *Nature Rev. Mol. Cell Biol.* **8**, 394–404 (2007).
4. Merideth, M. A. *et al.* Phenotype and course of Hutchinson–Gilford progeria syndrome. *N. Engl. J. Med.* **358**, 592–604 (2008).

5. Davies, B. S., Fong, L. G., Yang, S. H., Coffinier, C. & Young, S. G. The posttranslational processing of prelamin A and disease. *Annu. Rev. Genomics Hum. Genet.* **10**, 153–174 (2009).
6. Olive, M. *et al.* Cardiovascular pathology in Hutchinson–Gilford progeria: correlation with the vascular pathology of aging. *Arterioscler. Thromb. Vasc. Biol.* doi:10.1161/ATVBAHA.110.209460 (26 August 2010).
7. Ragnauth, C. D. *et al.* Prelamin A acts to accelerate smooth muscle cell senescence and is a novel biomarker of human vascular aging. *Circulation* **121**, 2200–2210 (2010).
8. Varga, R. *et al.* Progressive vascular smooth muscle cell defects in a mouse model of Hutchinson–Gilford progeria syndrome. *Proc. Natl Acad. Sci. USA* **103**, 3250–3255 (2006).
9. Pegoraro, G. *et al.* Ageing-related chromatin defects through loss of the NURD complex. *Nature Cell Biol.* **11**, 1261–1267 (2009).
10. Scaffidi, P. & Misteli, T. Reversal of the cellular phenotype in the premature aging disease Hutchinson–Gilford progeria syndrome. *Nature Med.* **11**, 440–445 (2005).
11. Dechat, T. *et al.* Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Genes Dev.* **22**, 832–853 (2008).
12. Scaffidi, P. & Misteli, T. Lamin A-dependent nuclear defects in human aging. *Science* **312**, 1059–1063 (2006).
13. McClintock, D. *et al.* The mutant form of lamin A that causes Hutchinson–Gilford progeria is a biomarker of cellular aging in human skin. *PLoS ONE* **2**, e1269 (2007).
14. Constantinescu, D., Gray, H. L., Sammak, P. J., Schatten, G. P. & Csoka, A. B. Lamin A/C expression is a marker of mouse and human embryonic stem cell differentiation. *Stem Cells* **24**, 177–185 (2006).
15. Freberg, C. T., Dahl, J. A., Timoskainen, S. & Collas, P. Epigenetic reprogramming of OCT4 and NANOG regulatory regions by embryonal carcinoma cell extract. *Mol. Biol. Cell* **18**, 1543–1553 (2007).
16. Peric-Hupkes, D. *et al.* Molecular maps of the reorganization of genome–nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603–613 (2010).
17. Irizarry, R. A. *et al.* The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genet.* **41**, 178–186 (2009).
18. Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44–57 (2009).
19. Dennis, G. *et al.* DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* **4**, R60 (2003).
20. McClintock, D., Gordon, L. B. & Djabali, K. Hutchinson–Gilford progeria mutant lamin A primarily targets human vascular cells as detected by an anti-Lamin A G608G antibody. *Proc. Natl Acad. Sci. USA* **103**, 2154–2159 (2006).
21. Gorenne, I., Kavurma, M., Scott, S. & Bennett, M. Vascular smooth muscle cell senescence in atherosclerosis. *Cardiovasc. Res.* **72**, 9–17 (2006).
22. Minamino, T. & Komuro, I. Vascular cell senescence: contribution to atherosclerosis. *Circ. Res.* **100**, 15–26 (2007).
23. Huang, S. *et al.* Correction of cellular phenotypes of Hutchinson–Gilford progeria cells by RNA interference. *Hum. Genet.* **118**, 444–450 (2005).
24. Liu, G. H. *et al.* Regulation of myoblast differentiation by the nuclear envelope protein NET39. *Mol. Cell. Biol.* **29**, 5800–5812 (2009).
25. Washburn, M. P., Wolters, D. & Yates, J. R., III. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnol.* **19**, 242–247 (2001).
26. Ruis, B. L., Fattah, K. R. & Hendrickson, E. A. The catalytic subunit of DNA-dependent protein kinase regulates proliferation, telomere length, and genomic stability in human somatic cells. *Mol. Cell. Biol.* **28**, 6182–6195 (2008).
27. Li, H., Vogel, H., Holcomb, V. B., Gu, Y. & Hasty, P. Deletion of Ku70, Ku80, or both causes early aging without substantially increased cancer. *Mol. Cell. Biol.* **27**, 8205–8214 (2007).
28. Espejel, S. *et al.* Shorter telomeres, accelerated ageing and increased lymphoma in DNA-PKcs-deficient mice. *EMBO Rep.* **5**, 503–509 (2004).
29. Han, X. *et al.* Tethering by lamin A stabilizes and targets the ING1 tumour suppressor. *Nature Cell Biol.* **10**, 1333–1340 (2008).
30. Zhang, J. *et al.* A human iPSC model of Hutchinson–Gilford progeria reveals vascular smooth muscle and mesenchymal stem cell defects. *Cell Stem Cell* **8**, 31–45 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Comai for providing experimental material, M. Hetzer, J. Karlseder, J.-F. Deleuze, M. J. Barrero, C. Rodriguez Esteban and L. Gerace for helpful discussions, M. Marti for teratoma analysis, M. C. Llach for karyotyping, T. Berggren, M. Lutz, I. Dubova, S. Stewart, R. Dev, M. Li, L. Laricchia-Robbio, A. M. Goebel and J. Kim for technical help, and M. Schwarz for administrative help. G.-H.L. and L.K. were partially supported by a CIRM grant (TG2-01158). J.Q. was partially supported by an AFAR/ Ellison Medical Foundation postdoctoral fellowship; A.D.P. was partially supported by a NIH training grant T32 CA009370. This study was supported by grants from NIH R01-DA025779 (K.Z.), and NIH P41 RR011823 (J.Y.); the G. Harold and Leila Y. Mathers Charitable Foundation, Sanofi-Aventis, Ellison Medical Foundation, MICINN and Fundacion Cellex (JCIB).

Author Contributions G.-H.L. and J.C.I.B. conceived the experiments; G.-H.L., B.Z.B., S.R., D.D., J.Q., S.-L.Y., A.D.P., K.S., L.K., C.W., J.T. and H.L.F. performed the experiments and analysed the data; S.B., I.S.-M., K.Z., J.Y. and J.C.I.B. analysed the data; G.-H.L., S.R., B.Z.B., A.D.P., K.Z. and J.C.I.B. wrote the manuscript.

Author Information Microarray data have been deposited in NCBI-GEO under the accession number GSE24487. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.C.I.B. (belmonte@salk.edu, izpisua@cmrbeu).

METHODS

Cell culture. H9 hESCs (WiCell Research) and iPSCs were maintained on a layer of mitotically inactivated mouse embryonic fibroblasts (MEFs) in hESC medium: DMEM/F12 (Invitrogen) supplemented with 0.1 mM non-essential amino acids (Invitrogen), 1 mM GlutaMAX (Invitrogen), 20% Knockout Serum Replacement (Invitrogen), 55 μ M β -mercaptoethanol (Invitrogen) and 10 ng ml⁻¹ bFGF (Joint Protein Central). hESCs and iPSCs were also cultured in Matrigel (BD Biosciences) with mTeSR medium (Stem Cell Technologies). Human HGPS fibroblasts AG01972, AG11498, AG06297, and normal fibroblasts GM00038 (9 years), AG05247 (87 years), and AG09602 (92 years) were purchased from Coriell Cell Repository. BJ normal human fibroblasts (CRL-2522) were purchased from ATCC. All human fibroblasts were cultured at 37 °C in DMEM containing GlutaMAX, non-essential amino acids, sodium pyruvate, and 15% fetal bovine serum (FBS). Human aortic smooth muscle cells were purchased from Lonza and maintained in SmGM-2 medium (Lonza, Cat. # CC-3182).

Reagents. Antibodies were obtained from the following sources. Abcam: anti-NANOG (ab21624), anti-H3K9me3 (ab8898), anti-progerin (ab66587), anti-emerin (ab14208); anti-Ku70 (ab2172); Santa Cruz Biotechnology: anti-Oct-3/4 (sc-5279), anti-SOX2 (sc-17320), anti-HDAC1 (sc-7872), anti-DNAPKcs (sc-9051), anti-lamin A/C (sc-6215), anti-lamin A/C (sc-7293), anti-lamin B1 (sc-6217); Cell Signaling: anti-HP1 α (2616); anti-Ku80 (2753); R&D systems: anti-Foxa2 (AF2400); Millipore: anti-TRA-1-60 (MAB4360); Sigma: anti- β -Tubulin III (T2200), anti-SMA (A5228), anti-Flag (M2), and anti-tubulin (T5168); Dako: anti-calponin (clone CALP); anti-endoglin (clone SN6h); BD Transduction Laboratories: anti-LAP2 β (611000); MBL: agarose-conjugated anti-GFP.

Plasmids. The pMXs vector containing the human cDNAs for *OCT4*, *SOX2*, *KLF4* and *c-MYC* were purchased from Addgene (17217, 17218, 17219 and 17220, respectively). pBABE-puro-GFP-progerin and pBABE-puro-GFP-wt-lamin A were purchased from Addgene (17663 and 17662, respectively). Flag-progerin lentiviral vector was kindly provided by L. Comai³¹. For the generation of the shRNA expression vectors against progerin and DNAPKcs, corresponding oligonucleotides (see Supplementary Table 5) were cloned into a MluI-ClaI-cleaved pLVTHM plasmid (Addgene, 12247). All the constructs generated were subjected to DNA sequencing to confirm accurate shRNA target sequence.

Retrovirus and lentivirus production. For retrovirus production, 293T cells were transfected with the pMXs vectors carrying *OCT4*, *SOX2*, *c-MYC*, *KLF4* or *GFP* cDNAs, together with the packaging plasmids (pCMV-gag-pol-PA and pCMV-VSVg, provided by G. Pao, The Salk Institute) using Lipofectamine 2000 (Invitrogen). Retroviruses were collected 36–48 h after transfection, and filtered through a 0.45- μ m filter. Lentiviruses were generated by co-transfecting the pLVTHM vector together with the packaging plasmids (psPAX2 and pMD2.G, from Addgene, 12260 and 12259 respectively) into 293T cells using Lipofectamine 2000 (Invitrogen). Lentiviruses were collected 36 h after transfection and concentrated by ultracentrifugation.

iPSCs generation. For the generation of human iPSCs, human fibroblasts were seeded in a 6-well plate and spin-infected with a mix of high-quality retroviruses encoding *OCT4*, *SOX2*, *KLF4*, *c-MYC* and *GFP* in the presence of 4 μ g ml⁻¹ polybrene. Three infections on consecutive days were performed. Six days after the first infection, fibroblasts were gently individualized with TrypLE (Invitrogen) and seeded onto fresh MEFs in the fibroblast culture medium. After 24 h, the medium was switched to hESC medium, and changed every 1–2 days depending on cell density. To establish the iPSC lines, colonies were manually picked and transferred onto MEF feeder cells for several passages before being transferred to Matrigel/mTeSR conditions.

Lentiviral infection of iPSCs. HGPS-iPSC#4 cell line cultured on Matrigel was treated with 10 μ M ROCK inhibitor Y-27632 for 1 h and then individualized with TrypLE. Cells were infected in suspension with either the concentrated lentivirus pLVTHM or pLVTHM-shRNA-progerin in the presence of ROCK inhibitor and polybrene (4 μ g ml⁻¹) for 1 h. Cells were centrifuged to remove the lentivirus and seeded back on fresh feeder MEFs in hESC media containing ROCK inhibitor. After growing for a few days, small colonies were manually passaged as a pool of colonies onto fresh MEFs to establish new iPSC lines. GFP expression was used as an indicator to determine successful integration of the lentiviruses.

Cell differentiation. For embryoid bodies (EBs) based differentiation, the iPSC colonies growing on MEFs were detached with dispase treatment, resuspended in DMEM/F12 medium supplemented with 0.1 mM non-essential amino acids, 0.5 mM L-glutamine, 10% FBS (Atlanta Biologicals), and 55 μ M β -mercaptoethanol and cultured in low attachment 6-well plates for 4 days. The EBs were then plated on gelatin-coated plates and maintained for another 10–17 days. Differentiation of iPSCs into fibroblasts was performed as previously described³². Directed differentiation towards smooth muscle cells (SMCs) was performed essentially as previously described³³ with slight modifications. Irrigated OP9 cells were plated at 1 \times 10⁵ cells per well onto gelatinized 6-well plates in OP9 growth medium. After the formation

of confluent iPSC cell cultures for 4 and 5 days, undifferentiated iPSC cells were harvested by treatment with 1 mg ml⁻¹ dispase and dispersed by scraping to maintain the cells in small clumps. Concurrently, iPSC cultures growing under the same conditions were used to obtain single cell suspension for counting. The iPSCs were added to OP9 cultures at a density of 3 \times 10⁵ per 2 ml per well of a 6-well plate in half TeSR medium and half hESC media. iPSCs were allowed to recover for 1–2 days in hESC media. At day 0 of differentiation, the media was changed to Knockout DMEM supplemented with 10% FBS (HyClone), 10 mM β -mercaptoethanol, 1 mM L-glutamine, and 100 mM nonessential amino acids. The iPSC/OP9 co-cultures were incubated for up to 10 days at 37 °C in 5% CO₂ conditions with medium change every other day. After 10 days of differentiation, the co-culture cells were harvested with TrypLE (Invitrogen) for single-cell suspension and labelled with CD34 microbeads kit (Miltenyi Biotec, Cat. #130-046-702). Following the manufacturer's protocol, cells were passed through MS separation column attached to a Midi-MACS separation unit (Miltenyi Biotec) to obtain a magnet-retained fraction of purified CD34⁺ cells. Isolated CD34⁺ cells were then plated in smooth muscle cell media (SmGM-2 BulletKit, Lonza, Cat. # CC-3182) and maintained at 37 °C in 5% CO₂ conditions with medium change every 2–3 days³⁴. SMCs were passaged using TrypLE (diluted 1:4) for 3 min at 37 °C. To analyse early onset of senescence, cells were passaged at a ratio of 1:3 (~6,000–7,500 cells per cm²) only when the cells reached confluence. To calculate population doublings, SMCs seeded at 3,500 cells per cm² were passaged once culture reached 85–90% confluence. Cell growth was measured at every passage by calculation of accumulated population doublings using the formula (logH – logS)/log2.0 where H = number of cells harvested, S = number of cells seeded on the first day of each passage.

Protein and mRNA analysis. Cells were lysed and subjected to immunoblotting analysis according to the previously described method³⁵. Total RNA was extracted using TRIzol (Invitrogen) followed by cDNA synthesis using High capability RNA-to-cDNA Mater Mix (Invitrogen). Quantitative RT-PCR was performed using SYBR Green PCR Master Mix (Applied Biosystems). Primer sequences are listed in Supplementary Table 5.

Immunofluorescence microscopy. Cells were fixed with 4% formaldehyde in PBS at room temperature for 20–30 min. After fixation, cells were treated with 0.4% Triton X-100 in PBS for 5 min at room temperature. After blocked with 10% FBS in PBS for 30 min, cells were incubated at room temperature for 1 h or at 4 °C overnight with the primary antibody, followed by washing in PBS and incubation at room temperature for 1 h with the corresponding secondary antibody. Nuclei were stained with Hoechst 33342 (Invitrogen). Quantitative microscopy measurements were carried out as described previously⁹. Error bars represent standard deviations.

Immunohistochemical detection of NANOG. Cells were fixed with 4% formaldehyde in PBS at room temperature for 30 min, and permeabilized with 0.4% Triton-X100 in PBS for 10 min. Then the cells were incubated overnight with rabbit anti-human NANOG antibody in 1% BSA/PBS, followed by incubation with a secondary biotin-conjugated anti-rabbit antibody for 2 h. Finally, cells were incubated with streptavidin-HRP for 1 h (Vector), and NANOG-positive cells were visualized with a DAB substrate kit (Vector).

Teratoma analysis. To test pluripotency *in vivo*, NOD-SCID IL2Rgamma^{null} mice (Jackson laboratories) were injected with the indicated iPSC lines and teratoma formation assessed. Briefly, ~10⁶ iPSCs in ~50 μ l of hESC medium were injected into the testis or kidney capsule of anesthetized mice. Mice were monitored for teratoma formation and euthanized ~6–12 weeks after injection. Teratomas were harvested, processed and analysed by haematoxylin–eosin staining and immunostaining. All animal experiments were performed with approval of The Salk Institute Institutional Animal Care and Use Committee (IACUC).

Mutation validation. Primer sequences to amplify exon 11 of the *LMNA* gene are listed in Supplementary Table 5. PCR (final volume 50 μ l) reactions using 3 ng genomic DNA templates, 100 nM of the forward and reverse primers with 25 μ l Taq 2 \times Master Mix (NEB) was performed at 94 °C for 2 min, 34 cycles of 94 °C 30 s, 55.5 °C for 40 s, and 72 °C for 40 s, and finally 72 °C for 3 min. Products were purified with 0.9 \times volume of AMPure beads (Agencourt). Amplicons were sequenced by capillary Sanger sequencing (Genewiz). Results were visualized using an ABI Sequence Scanner.

Genome-wide DNA methylation analysis. Genomic DNA was extracted using ALLPrep DNA/RNA Mini kit (Qiagen). Bisulphite conversion and capture reaction was carried out on each sample (genomic DNA of fibroblasts, iPSCs, or hESCs). The detailed protocol for genomic DNA methylation has been described previously³⁶, and the detailed information for DNA methylation is presented in Supplementary Table 1.

Bisulphite sequencing of OCT4 and LMNA promoters. Bisulphite conversion was carried out using 2 μ g of purified genomic DNA using the Zymo EZ-DNA Methylation Gold Kit (Zymo Research) following the manufacturer's instructions. PCR was set up using previously published primers¹⁵. Cycling was terminated at 35

cycles. PCR products were purified using 2% Size-Select E-gel (Invitrogen) and reamplified for 10 cycles using Phusion HF enzyme (NEB). PCR products were cloned using Zero-blunt PCR Cloning kit (Invitrogen) and heat transfected to TOP10 *Escherichia coli* competent cells (Invitrogen). Individual colonies were selected and sent for single pass sequencing.

DNA microarray and bioinformatics analysis. The GeneChip microarray processing was performed by the Functional Genomics Core in the Institute for Research in Biomedicine (Barcelona, Spain) according to the manufacturer's protocols (Affymetrix). The amplification and labelling were processed as indicated in Nugen protocol with 25 ng starting RNA. For each sample, 3.75 μ g ssDNA were labelled and hybridized to the Affymetrix HG-U133 Plus 2.0 chips. Expression signals were scanned on an Affymetrix GeneChip Scanner (7.0 upgrade). The data extraction was done by the Affymetrix GCOS software v.1.4.

The statistical analysis of the data was performed using ArrayStar 3. Briefly, raw CEL files were imported together with gene annotation from NetAffix (from 11/13/2009) and after checking for top replication quality for each of the five pairs of samples ($R^2 > 0.99$), data was summarized at the gene level (20,765 genes) and the median was used for each gene and sample type. As both H9 hESCs and HGPS-iPSCs originate from female samples, and in order to remove any possible bias introduced by the X and Y chromosome-coded genes, we performed the same analysis with only autosome genes (19,884 genes). The result of the hierarchical clustering is very similar to the one using all genes and is shown in Supplementary Fig. 9b. In addition, a principal component analysis was performed on RMA-normalized probeset intensity values for autosomes using the prcomp function in R (<http://www.r-project.org/>) (the same figure including all genes gave highly similar results, data not shown). A figure illustrating the two first principal components is shown in Supplementary Fig. 9c. Differences between some of the samples is shown using scatter plot of RMA-normalized intensity values in Supplementary Fig. 9d.

Multidimensional protein identification technology (MudPIT) analysis of progerin-associated proteins. The immunoprecipitation for MudPIT assay was performed as previously described^{24,25}. In brief, HEK293T cells were transfected with GFP-progerin or GFP and maintained in culture for 48 h. After cells were lysed, the GFP-progerin, GFP, and their associated proteins were immunoadsorbed to anti-GFP agarose. The immunoprecipitates were then eluted with 8 M urea in 100 mM Tris, pH 8.5. The samples were reduced by adding 0.3 μ l of 1 M TCEP (for a final concentration of 5 mM TCEP) and incubated at room temperature. To alkylate, 1.2 μ l of iodoacetamide (10 mM final concentration) was added and the samples were subsequently incubated at room temperature in the dark for 15 min. The addition of 180 μ l of 100 mM Tris pH 8.5 diluted the solutions to 2 M urea. Calcium chloride (100 mM) was then added (2.4 μ l) for a final concentration of 1 mM CaCl_2 . Trypsin (0.5 μ g μ l⁻¹) was added in the amount of 7.0 μ l. The resulting mixtures were then shaken for 18 h and incubated in the dark at 37 °C. To neutralize 13.5 μ l of formic acid (90%) was added for a final concentration of 5% formic acid. The tubes were centrifuged for 30 min at 2 °C in a table-top centrifuge. Upon completion of the digestion, the proteins were pressure-loaded onto a fused silica capillary desalting column containing 3 cm of 5- μ m strong cation exchange (SCX) followed by 3 cm of 5- μ m C18 (reverse phase or RP material) packed into an undeactivated 250- μ m i.d. capillary. Using 1.5 ml of buffer A (95% water, 5% acetonitrile and 0.1% formic acid) the desalting columns were washed overnight. Following the desalting process, a 100- μ m internal diameter capillary consisting of a 10- μ m laser pulled tip packed with 10 cm 3- μ m Aqua C18 material (Phenomenex) was attached to the filter union (desalting column-filter union-analytical column) and the entire split-column (desalting column-filter union-analytical column) was placed in line with an Agilent 1100 quaternary HPLC (Palo Alto, CA) and analysed using a modified 6-step separation, described previously²⁵. The buffer solutions used were 5% acetonitrile/0.1% formic acid (buffer A), 80% acetonitrile/0.1% formic acid (buffer B), and 500 mM ammonium acetate/5% acetonitrile/0.1% formic acid (buffer C). Step 1 consisted of a 90 min gradient from 0–100% buffer B. Steps 2–5 had the following profile: 3 min of 100% buffer A, 2 min of X% buffer C, a 10 min gradient from 0–15% buffer B, and a 97 min gradient from 15–45% buffer B. The 2 min buffer C percentages (X) were 20, 40, 60, 80% respectively for the 6-step analysis. In the final step, the gradient contained: 3 min of 100% buffer A, 20 min of 100% buffer C, a 10 min gradient from 0–15% buffer B, and a 107 min gradient from 15–70% buffer B. As peptides eluted from the microcapillary column, they were electrosprayed directly into an LTQ 2-dimensional ion trap mass spectrometer (ThermoFinnigan) with the application of a distal 2.4 kV spray voltage. A cycle of one full-scan mass spectrum (400–1,400 m/z) followed by 8 data-dependent MS/MS spectra at a 35% normalized collision energy was repeated continuously throughout each step of the multidimensional separation. Application of mass

spectrometer scan functions and HPLC solvent gradients were controlled by the Xcalibur data system.

As each step was executed, its spectra were recorded to a RAW file. This data was then converted into .ms2 format through the use of RawXtract (Version 1.9). From the .ms2 files, poor quality spectra were removed from the data set using an automated spectral quality assessment algorithm³⁷. MS/MS spectra remaining after filtering were searched with the SEQUEST algorithm³⁸ against the NCBI RefSeq Human (04-23-2010) protein database concatenated to a decoy database in which the sequence for each entry in the original database was reversed^{25,39}. All searches were parallelized and performed on a Beowulf computer cluster consisting of 100 1.2 GHz Athlon CPUs⁴⁰. No enzyme specificity was considered for any search. SEQUEST results were assembled and filtered using the DTASelect (version 2.0) program. DTASelect 2.0 uses a linear discriminant analysis to dynamically set XCorr and DeltaCN thresholds for the entire data set to achieve a user-specified false positive rate. The false positive rates are estimated by the program from the number and quality of spectral matches to the decoy database. The hits detected uniquely in the GFP-progerin sample but not in GFP sample represent proteins that are specifically associated with progerin, by either direct or indirect interactions.

Co-immunoprecipitation. BJ human fibroblasts were transduced with retrovirus encoding GFP-progerin, GFP-lamin A or GFP, and maintained in culture for 72 h. For immunoprecipitation, cells were lysed in ice-cold lysis buffer (250 mM NaCl, 0.5% Triton X-100, 50 mM Tris, pH 7.5, 1 mM EGTA, 1 mM EDTA, 10% glycerol, and complete protease inhibitor cocktail (Roche Diagnostics)). Samples were briefly sonicated and immunoprecipitated by incubating with anti-GFP agarose. The immunoprecipitates were washed extensively in lysis buffer, eluted in SDS sample buffer, and subjected to immunoblotting.

Senescence-associated beta-galactosidase (SA- β gal) assay. SA- β gal assay was performed based on previously described methods⁴¹.

Measurement of telomere length. Genomic DNA was isolated from 1×10^6 cells. The telomere-specific oligonucleotide probe (5'-TTAGGGTTAGGGTTAGGGTTAGGG-3'; ValueGene) was end-labelled using γ -³²P-ATP (MP Biomedicals) and T4 polynucleotide kinase (NEB). Two μ g of genomic DNA for each sample was digested with AluI (NEB) and MboI (NEB) and subjected to Southern analysis with the telomere-specific probe. Mean telomere length was calculated from $\sum OD_i / (\sum OD_i \sum MW_i)$. OD_i and MW_i are optical density and molecular weight at a given position i , respectively.

Cell proliferation assay. Cell proliferation was determined with CellTiter 96 AQueous One Solution Cell Proliferation Assay (MTS (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium)), according to the protocol provided by the manufacturer (Promega).

Statistical analysis. Results are presented as mean \pm s.d. Comparisons were performed with student's t -test or one-way anova. $P < 0.05$ was defined as statistically significant.

- Candelario, J., Borrego, S., Reddy, S. & Comai, L. Accumulation of distinct prelamin A variants in human diploid fibroblasts differentially affects cell homeostasis. *Exp. Cell Res.* **317**, 319–329 (2011).
- Hockemeyer, D. *et al.* A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell* **3**, 346–353 (2008).
- Vodyanik, M. A., Bork, J. A., Thomson, J. A. & Slukvin, I. I. Human embryonic stem cell-derived CD34+ cells: efficient production in the coculture with OP9 stromal cells and analysis of lymphohematopoietic potential. *Blood* **105**, 617–626 (2005).
- Lu, S. J., Ivanova, Y., Feng, Q., Luo, C. & Lanza, R. Hemangioblasts from human embryonic stem cells generate multilayered blood vessels with functional smooth muscle cells. *Regen. Med.* **4**, 37–47 (2009).
- Kawamura, T. *et al.* Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* **460**, 1140–1144 (2009).
- Deng, J. *et al.* Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nature Biotechnol.* **27**, 353–360 (2009).
- Bern, M., Goldberg, D., McDonald, W. H. & Yates, J. R., III. Automatic quality assessment of peptide tandem mass spectra. *Bioinformatics* **20** (Suppl 1), i49–i54 (2004).
- Eng, J., McCormack, A. & Yates, J. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976–989 (1994).
- Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J. & Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2**, 43–50 (2003).
- Sadygov, R. G. *et al.* Code developments to improve the efficiency of automated MS/MS spectra interpretation. *J. Proteome Res.* **1**, 211–215 (2002).
- Debacq-Chainiaux, F., Erusalimsky, J. D., Campisi, J. & Toussaint, O. Protocols to detect senescence-associated beta-galactosidase (SA- β gal) activity, a biomarker of senescent cells in culture and *in vivo*. *Nature Protocols* **4**, 1798–1806 (2009).

Duplications of the neuropeptide receptor gene *VIPR2* confer significant risk for schizophrenia

Vladimir Vacic^{1,2}, Shane McCarthy¹, Dheeraj Malhotra^{1,3,4}, Fiona Murray^{5,6}, Hsun-Hua Chou^{3,4}, Aine Peoples⁷, Vladimir Makarov^{8,9}, Seungtae Yoon^{8,9}, Abhishek Bhandari^{1,3,4}, Roser Corominas⁴, Lilia M. Iakoucheva⁴, Olga Krastoshevsky¹⁰, Verena Krause¹⁰, Verónica Larach-Walters¹¹, David K. Welsh^{4,12,13}, David Craig¹⁴, John R. Kelsoe^{4,12,15}, Elliot S. Gershon¹⁶, Suzanne M. Leal¹⁷, Marie Dell Aquila^{5,18}, Derek W. Morris⁷, Michael Gill⁷, Aiden Corvin⁷, Paul A. Insel^{5,6}, Jon McClellan¹⁹, Mary-Claire King^{20,21}, Maria Karayiorgou²², Deborah L. Levy¹⁰, Lynn E. DeLisi²³ & Jonathan Sebat^{1,3,4,15,24}

Rare copy number variants (CNVs) have a prominent role in the aetiology of schizophrenia and other neuropsychiatric disorders¹. Substantial risk for schizophrenia is conferred by large (>500-kilobase) CNVs at several loci, including microdeletions at 1q21.1 (ref. 2), 3q29 (ref. 3), 15q13.3 (ref. 2) and 22q11.2 (ref. 4) and microduplication at 16p11.2 (ref. 5). However, these CNVs collectively account for a small fraction (2–4%) of cases, and the relevant genes and neurobiological mechanisms are not well understood. Here we performed a large two-stage genome-wide scan of rare CNVs and report the significant association of copy number gains at chromosome 7q36.3 with schizophrenia. Microduplications with variable breakpoints occurred within a 362-kilobase region and were detected in 29 of 8,290 (0.35%) patients versus 2 of 7,431 (0.03%) controls in the combined sample. All duplications overlapped or were located within 89 kilobases upstream of the vasoactive intestinal peptide receptor gene *VIPR2*. *VIPR2* transcription and cyclic-AMP signalling were significantly increased in cultured lymphocytes from patients with microduplications of 7q36.3. These findings implicate altered vasoactive

intestinal peptide signalling in the pathogenesis of schizophrenia and indicate the VPAC2 receptor as a potential target for the development of new antipsychotic drugs.

A majority of the rare CNVs that have been implicated in schizophrenia involve large (>500 kb) genomic regions where local segmental duplication architecture promotes frequent and nearly identical rearrangements by non-allelic homologous recombination (NAHR). Because of the high structural mutation rates at these loci, the strong phenotypic effects of the causal variants, and the excellent power of most array platforms to detect such large CNVs, these genomic hotspots were the first to be detected in studies of CNVs in schizophrenia. As most of the genome lacks the duplication architecture of the NAHR hotspots described earlier and because a variety of mutational mechanisms can give rise to structural rearrangements, causal variants in other regions of the genome may consist of CNVs that are individually rarer and smaller (<500 kb) than those arising at NAHR hotspots. For example, microdeletions of the gene *neurexin 1* (*NRXN1*), which are highly enriched in autism and schizophrenia^{6,7},

Table 1 | Significant association of four CNV regions with schizophrenia

Region (hg18)	Genes	Band	Type	Primary Cases	Primary Controls	Secondary Cases	Secondary Controls	Peak OR (CI)	Peak P-value	Permutation P-value
chr22: 19786712–19795854	<i>BCR</i>	22q11.2	Del.	2	0	22	0	14.21* (4.24, infinity)	2.4×10^{-6}	$<5.00 \times 10^{-6}$
chr7: 158731401–158810016	<i>VIPR2</i> **	7q36.3	Dup.	2	0	18	1	16.41 (3.11, infinity)	8.39×10^{-5}	4.00×10^{-5}
chr16: 29569647–30209382	28 genes	16p11.2	Dup.	4	0	18	1	16.14 (3.06, infinity)	0.000097	0.0001
chr15: 29694064–29705665	<i>OTUD7A</i>	15q13.3	Del.	2	0	16	1	14.94 (2.80, infinity)	0.00023	0.00016
chr7: 158448321–158605936	<i>VIPR2</i> , <i>BC042556</i>	7q36.3	Dup.	2	0	12	0	8.26* (2.36, infinity)	0.00086	0.0007
chr15: 28881608–28991107	<i>MTMR15</i>	15q13.3	Del.	2	0	16	1	14.94 (2.80, infinity)	0.00023	0.001
chr3: 196826549–196872080	<i>CR597873</i> , <i>SDHALP2</i>	3q29	Dup.	2	0	8	0	5.65* (1.56, infinity)	0.01	0.005
chr6: 162835583–162997592	<i>PARK2</i>	6q26	Dup.	2	0	6	0	4.41* (1.17, infinity)	0.03	0.044

Events, ORs and exact conditional (EC) P-values listed here correspond to the peak of association. Empirical P-values for the entire target region were then computed based on permutation of case and control labels. The minimal threshold for statistical significance after Bonferroni correction for the 114 loci tested was permutation $P < 4.4 \times 10^{-4}$. When the number of controls in the secondary sample was 0, Haldane correction (adding 0.5 to each cell in the table) was applied in order to get a finite OR (*). All genes overlapping with the target region are listed or the closest gene within 100 kb (**). Del., deletion; Dup., duplication.

¹Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. ²Department of Computer Science, Columbia University, New York, New York 10027, USA. ³Beyster Center for Genomics of Psychiatric Diseases, University of California, San Diego, La Jolla, California 92037, USA. ⁴Department of Psychiatry, University of California, San Diego, La Jolla, California 92037, USA. ⁵Department of Medicine, University of California, San Diego, La Jolla, California 92037, USA. ⁶Department of Pharmacology, University of California, San Diego, La Jolla, California 92037, USA. ⁷Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine and Department of Psychiatry, Trinity College Dublin, Dublin 2, Ireland. ⁸Seaver Autism Center, Mount Sinai School of Medicine, New York, New York 10029, USA. ⁹Department of Psychiatry, Mount Sinai School of Medicine, New York, New York 10029, USA. ¹⁰McLean Hospital, Belmont, Massachusetts 02478, USA. ¹¹Faculty of Medicine, Universidad Andrés Bello, Santiago, Chile. ¹²Veterans Affairs San Diego Healthcare System, San Diego, California 92161, USA. ¹³Center for Chronobiology, University of California, San Diego, La Jolla, California 92037, USA. ¹⁴Neurogenetics Division, Translational Genomics Research Institute, Phoenix, Arizona 85004, USA. ¹⁵Institute for Genomic Medicine, University of California, San Diego, La Jolla, California 92037, USA. ¹⁶Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, Chicago, Illinois 60637, USA. ¹⁷Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77025, USA. ¹⁸Division of Medical Genetics, University of California, San Diego, La Jolla, California 92037, USA. ¹⁹Department of Psychiatry, University of Washington, Seattle, Washington 98195, USA. ²⁰Department of Genome Sciences, University of Washington, Seattle, Washington 98195, USA. ²¹Department of Medicine, University of Washington, Seattle, Washington 98195, USA. ²²Department of Psychiatry, Columbia University, New York, New York 10027, USA. ²³Department of Psychiatry, Boston VA Healthcare System and Harvard Medical School, Brockton, Massachusetts 02301, USA. ²⁴Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California 92037, USA.

consist of overlapping deletions with non-recurrent breakpoints. *NRXN1* deletions are not flanked by segmental duplications, and may occur by different mutational mechanisms such as non-homologous end joining (NHEJ) or DNA-replication-mediated rearrangement.

To identify novel schizophrenia genes, we investigated copy number variation genome-wide using an approach that detects enrichment of multiple overlapping rare variants. Regions of interest were defined in a

primary sample of 802 patients and 742 controls as genomic segments containing CNVs in at least two cases and in no controls. This discovery step yielded 114 genomic regions of interest. In the secondary cohort of 7,488 patients and 6,689 controls, we assessed the association of these regions with schizophrenia (Supplementary Table 2). All CNVs overlapping each of the 114 regions of interest were collected, and CNV breakpoints falling within each region were used to partition the region

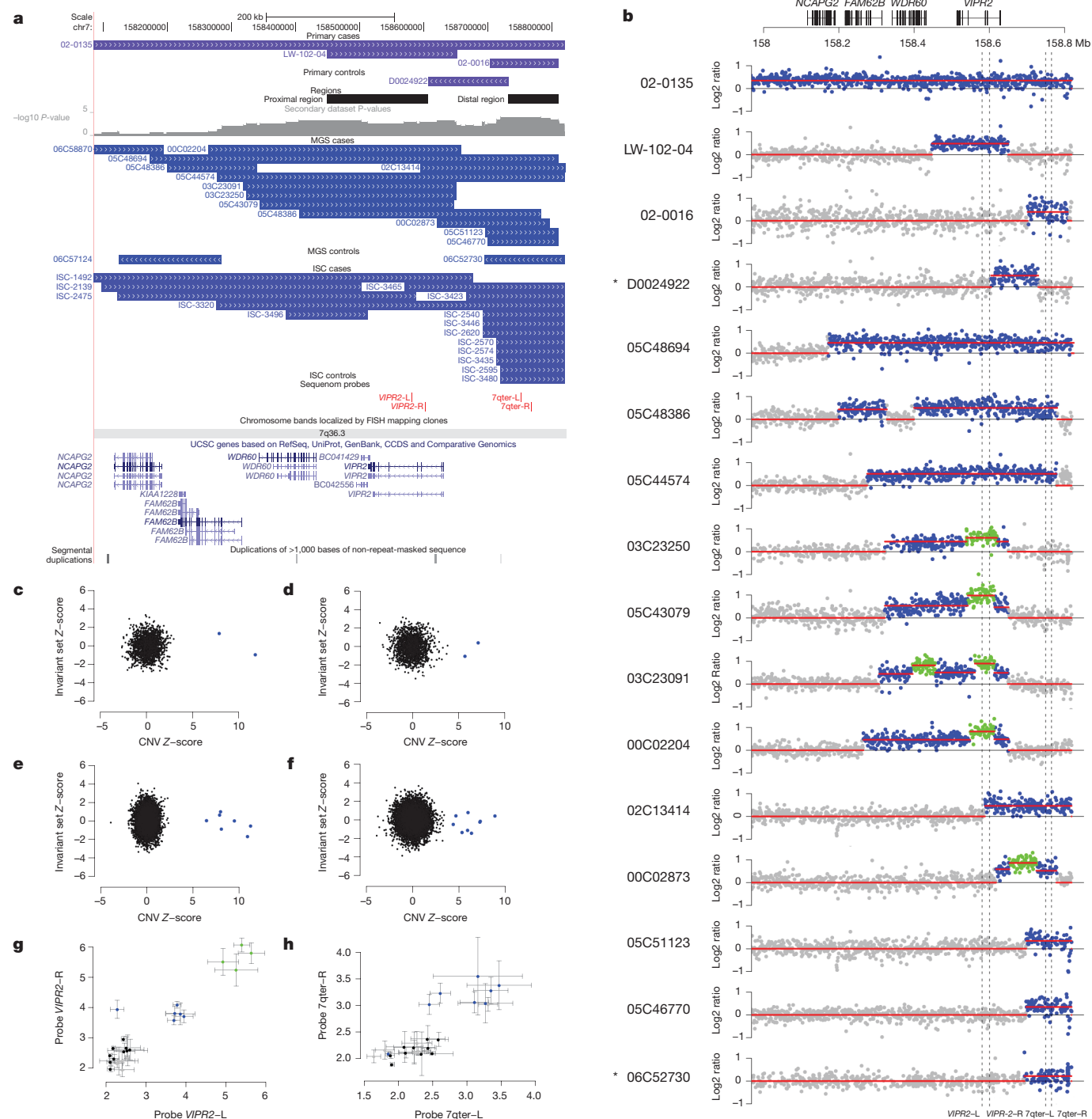


Figure 1 | Detection and validation of microduplications and triplications of 7q36.3. **a**, Map of CNVs detected in the primary and secondary cohorts from the UCSC genome browser (<http://genome.ucsc.edu>). ISC, International Schizophrenia Consortium. **b**, Plots of probe intensity ratios for 16 CNVs detected in the primary and MGS data sets. All are cases with the exception of two controls who are indicated with an asterisk. Regions with estimated copy numbers of 2, 3 and 4 are highlighted in grey, blue and green, respectively. Locations of four Sequenom validation assays are shown (dashed lines).

c–f, CNV genotypes were confirmed by MeZOD cluster plots of probe intensity ratios of the proximal and distal regions and in the primary data set (c and d, respectively) and secondary data set (e and f, respectively). g, h, Absolute copy numbers were confirmed for duplications and triplications of the proximal (*VIPR2*-L and *VIPR2*-R) and distal (7qter-L and 7qter-R) regions (b, dashed lines) by Sequenom MASSarray genotyping. Error bars represent the standard deviation of three replicates.

into a series of non-overlapping segments or bins (see Supplementary Fig. 1). Significance was tested within each bin by the exact conditional test, with ethnicity and study as covariates. The segment with the minimal *P*-value was defined as the peak of association within the region, and a permutation-based multiple testing correction scheme was applied to obtain the *P*-value for the region.

Of the 114 regions detected in the first step, four had statistically significant associations in the secondary sample after Bonferroni correction ($\alpha = 0.05/114 = 4.4 \times 10^{-4}$). Table 1 lists the four regions with significant *P*-values meeting this criterion and an additional four loci with nominally significant *P*-values ($P < 0.05$) in the secondary cohort. Regions with significant associations were loss of copy number at 22q11.2 ($P < 5 \times 10^{-6}$, odds ratio (OR) = 14.2), gain at 7q36.3 ($P = 4.0 \times 10^{-5}$, OR = 16.4), gain at 16p11.2 ($P = 1.0 \times 10^{-4}$, OR = 16.1) and loss at 15q13.3 ($P = 1.6 \times 10^{-4}$, OR = 14.9). No significant heterogeneity was observed for these genomic regions across studies (Breslow–Day–Tarone $P = 0.42 - 0.83$).

15q13.3, 16p11.2 and 22q11.2 are well-documented loci conferring increased risk for schizophrenia^{2,5,8}. All are hotspots for NAHR, and all alleles contributing to the association consist of large deletions with similar breakpoints. By contrast, microduplications at 7q36.3 have not been previously implicated in neuropsychiatric disorders. The 7q36.3 region harboured CNVs that overlapped but differed in size and breakpoint positions (Fig. 1a). The peak of association is located in the subtelomeric region of 7q, upstream of the gene *VIPR2*. Also, ranking fifth among the associations genome-wide was another region, 125 kb proximal to the peak at 7q36.3 ($P = 0.0007$; Table 1). Combining the two 7q36.3 regions into a single 362-kb region (chromosome 7 (chr7): 158448321–158810016), duplications were detected in 29 of 8,290 (0.35%) patients and 2 of 7,431 (0.03%) controls in this study. The *P*-value for the combined region in the combined sample was 5.7×10^{-7} and the odds ratio (OR) (95% confidence interval (CI)) was 14.1 (3.5, 123.9). A complete list of 7q36.3 CNVs is provided in Supplementary Table 3.

We examined sensitivity and specificity of CNV calls in the 7q36 region to determine the possibility of a spurious association (Supplementary Note). No additional duplications >100 kb were detected after

reducing the stringency of our CNV filtering criteria. Next, identical CNV calls were obtained using a more sensitive targeted CNV calling algorithm, median Z-score outlier detection (MeZOD)⁵ (Fig. 1c–f). All but one of the duplications (control sample 06C52730) were confirmed using the Sequenom MASSarray genotyping platform with assays designed for the proximal region (Fig. 1g) and for the distal region (Fig. 1h). Validated CNVs discovered in the Molecular Genetics of Schizophrenia (MGS) study subjects were mapped at higher resolution using the NimbleGen HD2 platform, and plots of probe intensity ratios from the HD2 array are shown in Fig. 1b and Supplementary Fig. 3. In addition, tandem duplications of the *VIPR2* gene were confirmed in two patients by fluorescence *in situ* hybridization (FISH) (Supplementary Fig. 4).

Unexpectedly, manual examination of probe ratios in Fig. 1b revealed additional structural complexity within some of the 7q36.3 CNVs. Copy number profiles in four patients (03C23250, 05C43079, 03C23091 and 00C02204) indicated that there were triplications nested within duplications of the proximal region (Fig. 1b). In all four patients, a triplication overlapped with exons 3 and 4 of the gene *VIPR2*. A copy number of four was confirmed in these samples using the Sequenom MASSarray CNV assay (Fig. 1g), and results for all samples were consistent with results in Fig. 1b. *VIPR2* transcripts were amplified from messenger RNA samples from the four triplication carriers. The normal *VIPR2* transcript was detected in all samples, and we did not observe a larger product corresponding to a transcript with duplicated exons.

Inheritance of the duplication at 7q36.3 could be evaluated in three families (Fig. 2). In family 02-135, the duplication was confirmed in the proband, but not detected in either of the unaffected parents, and thus is apparently *de novo* (Fig. 2f). In family 02-016, the duplication was detected in the proband and in a mother with a diagnosis of depression (Fig. 2d). In family LW102, the duplication was detected in the proband and in an unaffected mother. The proband's mother also had a son with a diagnosis of schizophrenia (LW-102-03) from a second marriage, but DNA was not available on this individual. Clinical psychiatric reports of patients 02-016 and 02-135 are provided in the Supplementary Note.

Variable expressivity is often characteristic of pathogenic CNVs⁵. We evaluated the spectrum of psychiatric phenotypes associated with

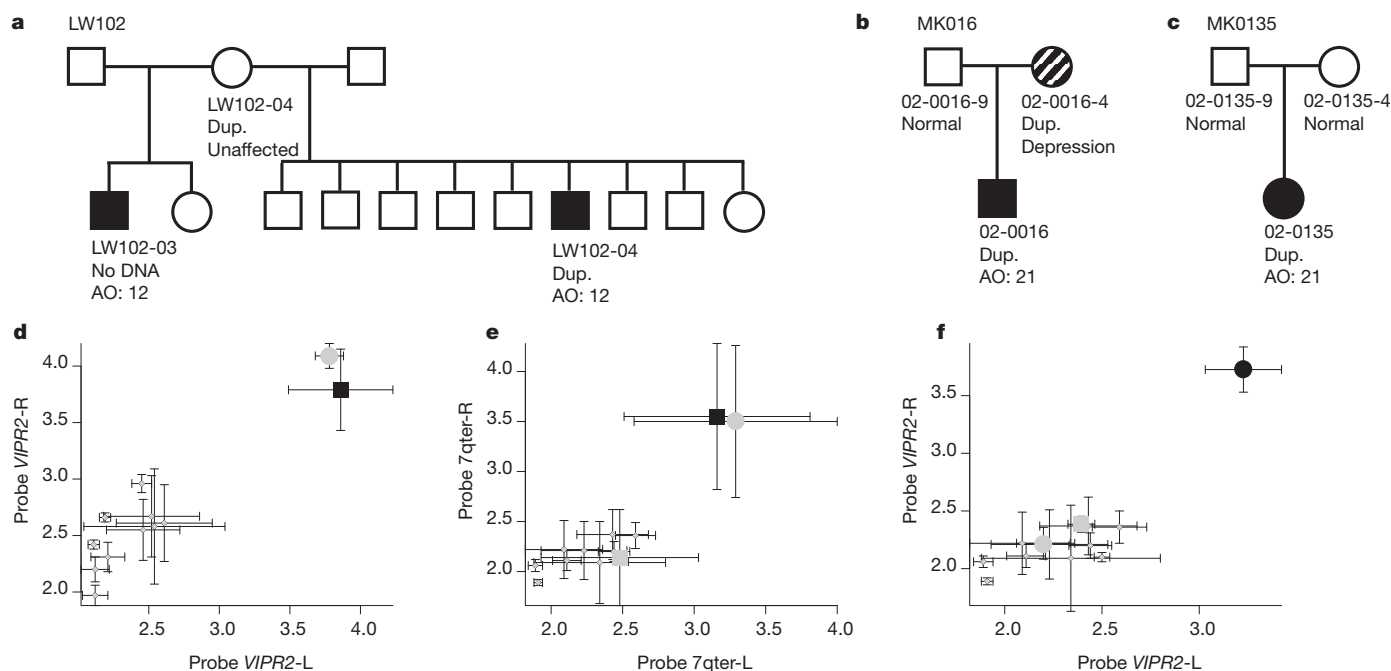


Figure 2 | Patterns of CNV inheritance in families. a–f, Pedigree diagrams are shown for families LW102 (a), 02-016 (b) and 02-135 (c), along with the Sequenom validation for families LW102 (d), 02-016 (e) and 02-135 (f). Sequenom validation was performed on a mother and one of the affected

sons (a), and all three family members (b), along with 10 CEU HapMap controls. Sequenom assays confirmed that duplications (Dup.) were present in the patients and maternally inherited from LW102-2 and 02-0016-4. AO, age at onset. Error bars represent the standard deviation of three replicates.

7q36.3 duplications by screening for these events in individuals with bipolar disorder or autism spectrum disorder (ASD). Microarray data were available for 2,777 patients from the Bipolar Genome Study (BiGS), for 996 ASD patients from the Autism Genome Project Consortium (AGP), and from our unpublished analyses of 114 patients with ASD using the NimbleGen HD2 platform. Microduplications of 7q36.3 (>100 kb in size) were detected in 3 of 1,110 (0.27%) of patients with ASD; compared with the controls described earlier, $P = 0.018$. Microduplications at 7q36.3 were detected in 2 of 2,777 (0.07%) patients with bipolar disorder; compared with the controls, $P = 0.23$. These results provide preliminary evidence that the clinical phenotypes associated with 7q36.3 duplications may include paediatric neurodevelopmental disorders such as autism, but do not include bipolar disorder. Also worthy of note, larger chromosomal abnormalities involving 7q have been reported in association with neurodevelopmental disorders, including deletions of 7q36-7qter^{9,10} and duplications of 7q35-7qter¹¹; a 550 kb duplication of 7qter (of unknown clinical relevance) has been reported in a patient with neurofibromatosis¹².

These genetic data implicate the gene *VIPR2*. All variants contributing to the observed association at 7q36.3 overlap with this gene or lie within the geneless subtelomeric region <89 kb from the transcriptional start site of *VIPR2*. *VIPR2* encodes the vasoactive intestinal peptide (VIP) receptor VPAC2, a G-protein-coupled receptor that is expressed in a variety of tissues including, in the brain, the suprachiasmatic nucleus, hippocampus, amygdala and hypothalamus¹³. VPAC2 binds VIP¹⁴, activates cyclic AMP (cAMP)-signalling and PKA, regulates synaptic transmission in the hippocampus^{15,16}, and promotes the proliferation of neural progenitor cells in the dentate gyrus¹⁷. Genetic studies in mouse have established that VIP signalling has a role in learning and memory¹⁸. VPAC2 also has a role in sustaining normal circadian oscillations in the suprachiasmatic nucleus¹⁹, and *VIPR2*-null²⁰ and *VIPR2*-overexpression²¹ mice exhibit abnormal rhythms of rest and activity.

cAMP signalling has been implicated in schizophrenia^{22,23}. We proposed that increases in *VIPR2* transcription and VPAC2-mediated cAMP signalling would be a consequence of the microduplications at 7q36.3. We thus assessed *VIPR2* mRNA and cAMP accumulation in response to VIP and a VPAC2-selective agonist (BAY 55-9837) in lymphoblastoid cell lines from eight MGS study subjects: two with subtelomeric duplications, three with duplications of *VIPR2*, four with partial triplications, and four controls with normal copy number of the region (see Supplementary Note). *VIPR2* transcripts were present at low but measurable levels in all cell lines. *VIPR2* mRNA levels were significantly increased in duplication carriers compared with controls (Fig. 3a). Likewise, cAMP responses to VIP and BAY 55-9837 were significantly greater in lymphoblastoid lines from carriers as compared to controls (Fig. 3b). In contrast, we observed no group difference in cAMP accumulation in response to a different G-protein-coupled receptor agonist, prostaglandin E2, thus confirming that the effect of 7q36 duplications on cAMP accumulation is mediated by the VPAC2 receptor.

The expression patterns that we observe indicate that a variety of different genomic duplications can influence the transcription of *VIPR2*. The exact genetic mechanism for this is unclear. Given that some risk variants are upstream of the gene and others are complex rearrangements that could potentially disrupt the duplicate copy, our results cannot be explained simply by an increase in gene dosage. It is likely that duplications of 7q36 affect the regulation of *VIPR2*. Tandem duplication of regulatory sequences, for instance, could affect expression of the gene. Alternatively, the subtelomeric location of *VIPR2* could be relevant to the mechanism. Intrinsic regulation of telomere structure and function often affects the transcriptional regulation of adjacent genes, a phenomenon known as telomere position effect^{24,25}. If *VIPR2* is under such epigenetic regulation, any large tandem duplication of the subtelomeric region could potentially cause the gene to escape repression. However, further studies are needed to determine the mechanism by which structural variants influence *VIPR2* expression.

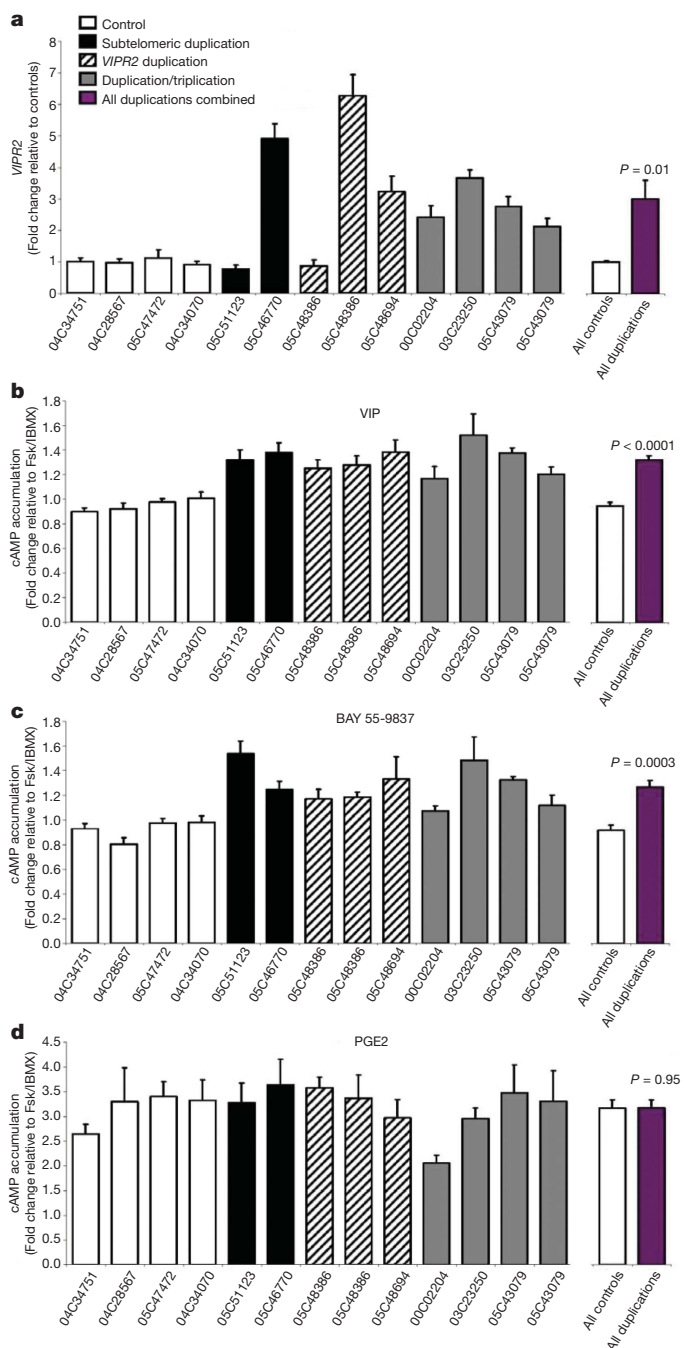


Figure 3 | Duplications and triplications of 7q36.3 result in increased *VIPR2* transcription and cAMP signalling. **a**, Quantitative PCR results of *VIPR2* mRNA from lymphoblastoid cell lines. Two to four subjects were tested for each of four genotypes (subtelomeric duplication, *VIPR2* duplication, exon 3/4 triplication and normal diploid copy number as control). Results are expressed as the mean fold-change of the sample relative to the mean of control samples. **b**, **c**, cAMP accumulation was measured in the same cell lines in response to VIP (100 nM) and the VPAC2 agonist BAY 55-9837 (100 nM). Results are expressed as fold-change over forskolin (Fsk)/IBMX (3-isobutyl-1-methylxanthine, a phosphodiesterase inhibitor) alone. **d**, No significant differences were observed in cAMP response to another G-protein-coupled receptor agonist, prostaglandin E2 (PGE2, 1 μ M), demonstrating that the effects are specific to VPAC2. For subjects, error bars represent standard error of the mean computed across replicates. Differences between the groups of nine duplication carriers and four controls were tested using unpaired two-tailed *t*-test.

In light of the emerging roles of *VIPR2* in the brain, our results support the hypothesis that the pathogenesis of schizophrenia—in some patients—involves the dysregulation of cellular processes such

as adult neurogenesis and synaptic transmission and of the corresponding cognitive processes of learning and memory. Furthermore, in light of the brain expression patterns of *VIPR2* (ref. 13), our results support the involvement of certain brain regions, such as the hippocampus, amygdala and the suprachiasmatic nucleus.

The link between *VIPR2* duplications and schizophrenia may have significant implications for the development of molecular diagnostics and treatments for this disorder. Genetic testing for duplications of the 7q36 region could enable the early detection of a subtype of patients characterized by overexpression of *VIPR2*. Significant potential also exists for the development of therapeutics targeting this receptor. For instance, a selective antagonist of the VPAC₂ receptor could have therapeutic potential in patients who carry duplications of the *VIPR2* region. Peptide derivatives²⁶ and small molecules²⁷ have been identified that are selective VPAC₂ inhibitors, and these pharmacological studies offer potential leads in the development of new drugs. Although duplications of *VIPR2* account for a small percentage of patients, the rapidly growing list of rare CNVs that are implicated in schizophrenia indicates that this psychiatric disorder is, in part, a constellation of multiple rare diseases¹. This knowledge, along with a growing interest in the development of drugs targeting rare disorders²⁸, provides an avenue for the development of new treatments for schizophrenia.

METHODS SUMMARY

Cohort description. Our primary cohort consisted of unrelated patients derived from family-based studies conducted by investigators at the University of Washington, McLean Hospital, Columbia University, Trinity College Dublin, New York University and Harvard Medical Schools (Supplementary Table 1). All samples were analysed by array comparative genomic hybridization (CGH) using the NimbleGen HD2 platform. The secondary cohort consisted of Affymetrix SNP Array 6.0 data from the MGS study of schizophrenia²⁹, publicly available data from the International Schizophrenia Consortium (ISC), genotyped using Affymetrix 6.0 and 5.0 platforms³, and Affymetrix 6.0 data on an independent set of controls from the BiGS³⁰ (Supplementary Table 1).

Intensity data processing and rare CNV calling. With the exception of the published CNV calls from the ISC, all data were processed and analysed centrally as follows. Microarray intensity data were normalized, and CNV calls were generated using an analysis package that we developed called C-score. All CNV call sets were filtered in a consistent fashion. To minimize the differential sensitivity of the various array platforms to detect CNVs, we limited our analysis to CNVs >100 kb. This size range is readily detectable by all platforms used in this study. The same criteria have been previously applied to filter CNVs across studies³. Last, sensitivity to detect large (>100 kb) copy number polymorphisms (CNPs) was evaluated at several locations in the genome. Overall sensitivity to detect CNVs was comparable in cases and controls in both cohorts. Additional details regarding C-score, statistical methods and evaluation of CNV call sets are described in the Supplementary Note.

Received 22 October 2010; accepted 28 January 2011.

Published online 23 February 2011.

- Sebat, J., Levy, D. L. & McCarthy, S. E. Rare structural variants in schizophrenia: one disorder, multiple mutations; one mutation, multiple disorders. *Trends Genet.* **25**, 528–535 (2009).
- The International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* **455**, 237–241 (2008).
- Mulle, J. G. *et al.* Microdeletions of 3q29 confer high risk for schizophrenia. *Am. J. Hum. Genet.* **87**, 229–236 (2010).
- Karayiorgou, M. *et al.* Schizophrenia susceptibility associated with interstitial deletions of chromosome 22q11. *Proc. Natl Acad. Sci. USA* **92**, 7612–7616 (1995).
- McCarthy, S. E. *et al.* Microduplications of 16p11.2 are associated with schizophrenia. *Nature Genet.* **41**, 1223–1227 (2009).
- Szatmari, P. *et al.* Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nature Genet.* **39**, 319–328 (2007).
- Rujescu, D. *et al.* Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum. Mol. Genet.* **18**, 988–986 (2008).
- Stefansson, H. *et al.* Large recurrent microdeletions associated with schizophrenia. *Nature* **455**, 232–236 (2008).
- Tyson, C. *et al.* Submicroscopic deletions and duplications in individuals with intellectual disability detected by array-CGH. *Am. J. Med. Genet. A* **139A**, 173–185 (2005).
- Wu, Y. *et al.* Submicroscopic subtelomeric aberrations in Chinese patients with unexplained developmental delay/mental retardation. *BMC Med. Genet.* **11**, 72 (2010).

- Morava, E. *et al.* Small inherited terminal duplication of 7q with hydrocephalus, cleft palate, joint contractures, and severe hypotonia. *Clin. Dysmorphol.* **12**, 123–127 (2003).
- Bartsch, O. *et al.* Two independent chromosomal rearrangements, a very small (550 kb) duplication of the 7q subtelomeric region and an atypical 17q11.2 (NF1) microdeletion, in a girl with neurofibromatosis. *Cytogenet. Genome Res.* **119**, 158–164 (2007).
- Sheward, W. J., Lutz, E. M. & Harmar, A. J. The distribution of vasoactive intestinal peptide 2 receptor messenger RNA in the rat brain and pituitary gland as assessed by *in situ* hybridization. *Neuroscience* **67**, 409–418 (1995).
- Fahrenkrug, J. Transmitter role of vasoactive intestinal peptide. *Pharmacol. Toxicol.* **72**, 354–363 (1993).
- Yang, K. *et al.* Vasoactive intestinal peptide acts via multiple signal pathways to regulate hippocampal NMDA receptors and synaptic transmission. *Hippocampus* **19**, 779–789 (2009).
- Waschek, J. A. Vasoactive intestinal peptide: an important trophic factor and developmental regulator? *Dev. Neurosci.* **17**, 1–7 (1995).
- Zaben, M. *et al.* The neurotransmitter VIP expands the pool of symmetrically dividing postnatal dentate gyrus precursors via VPAC₂ receptors or directs them toward a neuronal fate via VPAC₁ receptors. *Stem Cells* **27**, 2539–2551 (2009).
- Chaudhury, D., Loh, D. H., Dragich, J. M., Hagopian, A. & Colwell, C. S. Select cognitive deficits in vasoactive intestinal peptide deficient mice. *BMC Neurosci.* **9**, 63 (2008).
- Brown, T. M., Colwell, C. S., Waschek, J. A. & Piggins, H. D. Disrupted neuronal activity rhythms in the suprachiasmatic nuclei of vasoactive intestinal polypeptide-deficient mice. *J. Neurophysiol.* **97**, 2553–2558 (2007).
- Harman, A. J. *et al.* The VPAC₂ receptor is essential for circadian function in the mouse suprachiasmatic nuclei. *Cell* **109**, 497–508 (2002).
- Underhill, P. A. *et al.* Y chromosome sequence variation and the history of human populations. *Nature Genet.* **26**, 358–361 (2000).
- Millar, J. K. *et al.* DISC1 and PDE4B are interacting genetic factors in schizophrenia that regulate cAMP signaling. *Science* **310**, 1187–1191 (2005).
- Turetsky, B. I. & Moberg, P. J. An odor-specific threshold deficit implicates abnormal intracellular cyclic AMP signaling in schizophrenia. *Am. J. Psychiatry* **166**, 226–233 (2009).
- Gottschling, D. E., Aparicio, O. M., Billington, B. L. & Zakian, V. A. Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. *Cell* **63**, 751–762 (1990).
- Koering, C. E. *et al.* Human telomeric position effect is determined by chromosomal context and telomeric chromatin integrity. *EMBO Rep.* **3**, 1055–1061 (2002).
- Moreno, D. *et al.* Development of selective agonists and antagonists for the human vasoactive intestinal polypeptide VPAC₂ receptor. *Peptides* **21**, 1543–1549 (2000).
- Chu, A., Caldwell, J. S. & Chen, Y. A. Identification and characterization of a small molecule antagonist of human VPAC₂ receptor. *Mol. Pharmacol.* **77**, 95–101 (2009).
- Braun, M. M., Farag-El-Massah, S., Xu, K. & Cote, T. R. Emergence of orphan drugs in the United States: a quantitative assessment of the first 25 years. *Nature Rev. Drug Discov.* **9**, 519–522 (2010).
- Shi, J. *et al.* Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* **460**, 753–757 (2009).
- Zhang, D. *et al.* Singleton deletions throughout the genome increase risk of bipolar disorder. *Mol. Psychiatry* **14**, 376–380 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This study was supported by a gift from Ted and Vada Stanley to the Cold Spring Harbor Laboratory, a gift to J.S. from the Beyster family foundation, NIH grants to J.S. (MH076431, HG04222), D.L.L. (MH071523), M.-C.K. (MH083989), P.A.I. (GM66232), F.M. (HL091061), D.K.W. (MH082945), M.K. (MH061399), L.E.D. (MH044245), grants to J.S., D.K.W., D.L.L. and M.-C.K. from NARSAD, grants to A.C., M.G., D.M. from the Wellcome Trust (072894/Z/03/Z) and Science Foundation Ireland (08INIB1916), a career development award to D.K.W. by the Veterans Administration, and grants to D.L.L. from the Sidney R. Baer, Jr Foundation and Essel Foundation. We wish to thank the Genetic Association Information Network (GAIN), Molecular Genetics of Schizophrenia (MGS) and the Bipolar Genome Study (BiGS) for providing data for this study (investigators listed in the Supplementary Note). We wish to thank B. Trask, R. Malinow and J. Gleeson for discussions.

Author Contributions V.V. and J.S. wrote the manuscript. V.V., S.M., D.M., H.-H.C., F.M., V.M., S.Y., S.M.L., P.A.I. and J.S. designed the analytical strategy and analysed the data. M.G., A.C., J.M., M.K., D.L.L., V.L.-W. and L.E.D. oversaw the recruitment and clinical assessment of study participants. M.D.A. performed cytogenetic analysis. A.B., A.P. and D.M. designed genotype assays and performed genotyping. H.-H.C. and R.C. performed mRNA studies. F.M. performed biochemical studies. O.K., V.K., D.W.M., V.L.-W., L.E.D. and M.K. contributed to the interpretation of clinical patient data. J.M., M.-C.K., M.K., D.L.L., L.E.D., D.C., J.R.K. and E.S.G. contributed to the interpretation of data from genetic studies. P.A.I., L.M.I. and D.K.W. contributed to the interpretation of data from functional studies. J.S. coordinated the study.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.S. (jsebat@ucsd.edu).